Amy Webb

FACEBOOK

TENCENT

IBM

MICROSOFT

CEI NOUĂ TITANI TECH

Cum va schimba inteligența artificială cursul omenirii.

GOOGLE

AMAZON

BAYOU

APPLE

ALIBABA

globo
EDITURA

RECENZII DESPRE CEI NOUĂ TITANI TECH

„*Cei nouă titani tech* e o carte provocatoare, lizibilă și în care te regăsești cu ușurință. Amy Webb demonstrează o cunoaștere aprofundată a științei din spatele inteligenței artificiale și a tensiunilor geopolitice care ar putea apărea, mai ales între China și Statele Unite. Autoarea ne oferă o analiză profundă a felului în care inteligența artificială ne-ar putea recalibra economiile și actuala ordine mondială, oferind un plan amănunțit prin care omenirea ar putea să-și aleagă un viitor mai bun."

—Anja Manuel, Universitatea Stanford, co-fondator și partener la RiceHadleyGates

„*Cei nouă titani tech* e o carte importantă și cu idei proaspete care ne pune în perspectivă potențialul și pericolele inteligenței artificiale. Își va sluji inteligența artificială stăpânii din Washington, Silicon Valley și Wall Street, sau va sluji ea interesele publicului larg? Va reuși ea să concentreze sau să disperseze puterea economică și geopolitică? Grație lui Amy Webb putem înțelege întrebările și cum putem ajunge la răspunsurile care vor fi mai bune pentru omenire decât parcursul pe care ne aflăm deocamdată. *Cei nouă titani tech* ar trebui introdusă în sălile de clasă și de consiliu din lumea întreagă."

—Alec Ross, autorul *Industries of the Future*

„*Cei nouă titani tech* ne oferă perspective îndrăznețe referitor la viitorul inteligenței artificiale. Spre deosebire însă de mulți alții care încearcă să anticipeze viitorul, Webb nu se folosește de senzaționalism ci de argumente studiate, un context istoric profund și un nivel înfricoșător de plauzibilitate."

—Jonathan Zittrain, profesor de drept internațional și informatică, Universitatea Harvard

„*Cei nouă titani tech* ne pune pe gânduri, privind situația în perspectivă și mai ales punând întrebările corecte referitoarea la inteligența artificială și oferindu-ne o foaie de parcurs pentru un viitor optimist împreună cu inteligența artificială."

—Peter Schwartz, autorul cărții *The Art of the Long View*

„*Cei nouă titani tech* oferă argumente importante pentru a refuza abordările orientate spre prezent și a nu considera că firmele care dezvoltă inteligența artificială sunt morale prin definiție. Scenariile propuse de Webb pentru viitor sunt superbe, oferind o viziune amănunțită atât asupra aspectelor care trebuie evitate cât și a celor care trebuie îndeplinite."

—John C. Havens, director general IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, autor al cărții *Heartificial Intelligence: Embracing our Humanity to Maximize Machines*

CEI NOUĂ TITANI TECH

CEI NOUĂ TITANI TECH

Cum va schimba
inteligența artificială
cursul omenirii

AMY WEBB

globo
EDITURA

The original title of this book is:
The Big Nine: how the tech titans and their thinking machines could warp humanity
by Amy Webb

Copyright © 2019 by Amy Webb

Copyright © Editura Globo, 2019, pentru ediţia în limba română

Descrierea CIP a Bibliotecii Naţionale a României
WEBB, AMY
Cei 9 titani tech : cum va schimba inteligenţa artificială cursul omenirii / Amy Webb. - Bucureşti : Globo, 2019
ISBN 978-606-9653-00-5

004

Traducător: Răzvan Lerescu
Redactor: Diana Morărașu
Tehnoredactor: Sorin Sava

Comenzi:
telefon: 0726228911
www.edituraglobo.ro
comenzi@edituraglobo.ro

Tatălui meu, Don Webb,
omul cel mai inteligent pe care
l-am cunoscut vreodată.

CONȚINUT

Introducere: Înainte de a fi prea târziu 1

Partea I: Spiritul din aparat 15

1 Mintea și mașinile: o foarte scurtă istorie a inteligenței artificiale 17

2 Lumea insulară a triburilor inteligenței artificiale 66

3 O mie de tăieturi cu hârtia: consecințele neintenționate ale inteligenței artificiale 126

Partea a II-a: Viitorurile noastre 171

4 De la prezent la super-inteligența artificială: semnalele de avertizare 173

5 Prosperitate în a treia epocă a calculatoarelor: scenariul optimist 199

6 Cum să învățăm să trăim cu milioane de tăieturi cu hârtia: scenariul pragmatic 229

7 Dinastia Réngōng Zhìnéng: scenariul catastrofic 265

Partea a III-a: Rezolvarea problemelor 297

8 Pietricele și bolovani: cum să reparăm viitorul inteligenței artificiale 299

Mulțumiri 335

Bibliografie 339

Note 349

INTRODUCERE

ÎNAINTE DE A FI PREA TÂRZIU

Inteligența artificială face deja parte din viața noastră, însă nu sub forma pe care o așteptasem cu toții. magnifici

Inteligența artificială reprezintă structura discretă de rezistență a sistemelor noastre financiare, a rețelelor energetice și a lanțurilor de aprovizionare comerciale. Ea reprezintă infrastructura invizibilă care ne conduce prin trafic, găsește înțelesul corect în cuvintele noastre tastate greșit și stabilește ce ar trebui să cumpărăm, urmărim, ascultăm sau citim. Ea este tehnologia pe care se construiește viitorul nostru fiindcă se intersectează cu fiecare aspect din viața noastră: sănătatea și medicina, locuințele, agricultura, transporturile, sporturile și chiar iubirea, sexul și moartea.

Inteligența artificială nu e un trend în tehnologie, sau un cuvânt cheie, sau o distracție de moment. Ea reprezintă a treia eră a informaticii. Ne aflăm în mijlocul unor transformări importante, asemănătoare cu cele pe care le-au trăit generațiile din epoca revoluției industriale. La început, oamenii nu au putut să recunoască tranziția prin care treceau pentru că schimbarea s-a petrecut treptat, dacă ne raportăm la durata lor de viață. La finalul procesului, lumea arăta diferit: Marea Britanie și Statele Unite erau cele două puteri dominante în lume, având suficient capital industrial, militar și politic pentru a influența parcursul următorului secol.

Toată lumea dezbate până la saturație inteligența artificială și implicațiile acesteia pentru viitorul nostru. Cunoașteți deja argumentele cele mai frecvente: roboții ne vor lua locurile de muncă, roboții vor da economia peste cap, roboții vor sfârși prin a ucide oameni. Dacă înlocuim „mașină" cu „robot" vom regăsi aceleași dezbateri pe care le aveau oamenii acum 200 de ani. E normal să ne gândim asupra impactului noilor tehnologii asupra slujbelor noastre și a capacității noastre de a câștiga bani, deoarece vedem că multe industrii trec prin transformări profunde. E de înțeles de ce atunci când ne gândim la inteligența artificială, mintea noastră se gândește la exemple celebre precum HAL 9000 din filmul *2000: Odiseea spațială*, WOPR din *Simulări de război*, Skynet din *Terminator*, Rosie din *Familia Jetson*, Dolores din *Westworld* sau oricare dintre sutele de alte forme antropomorfizate ale inteligenței artificiale din cultura pop. Dacă nu lucrați direct în interiorul ecosistemului inteligenței artificiale, viitorul vi se poate părea fie extraordinar, fie îngrozitor, dar asta din niște motive complet greșite.

Persoanele neimplicate în cercetarea și dezvoltării inteligenței artificiale nu pot să vadă în mod clar semnalele, motiv pentru care dezbaterea publică pe această temă se rezumă la roboți care vor să domine lumea, sau reflectă un soi de optimism maniacal, îndrăzneț. Lipsa nuanțelor constituie una dintre problemele inteligenței artificiale încă de la începuturile acesteia: unii supraestimează dramatic domeniile în care ar putea fi aplicată, în vreme ce alții susțin că va deveni o armă de neoprit.

Știu asta pentru că mi-am petrecut o bună parte din ultimul deceniu în domeniul cercetării inteligenței artificiale și m-am întâlnit cu oameni și organizații atât din interiorul cât și din afara

ecosistemului inteligenței artificiale. Am fost consultant pentru diferite companii de vârf din domeniu, printre care se numără Microsoft sau IBM. M-am întâlnit și am consiliat diverse persoane cointeresate: investitori și manageri de fonduri mutuale, lideri din Departamentul Apărării și Departamentul de Stat, și diferiți parlamentari care cred că viitorul nu poate exista decât într-un cadru reglementat. Am avut și sute de întâlniri cu cercetători și ingineri care lucrează chiar în tranșee. Cei care lucrează direct cu inteligența artificială nu împărtășesc decât rareori viziunile extreme -fie utopice, fie apocaliptice- despre care auzim în presă.

Motivul e că cercetătorii din alte domenii ale științei, cei care chiar pun umărul la construirea viitorului inteligenței artificiale vor să tempereze așteptările. Depășirea unor limite istorice se face cu răbdare, timp, bani și statornicie- uităm constant acest aspect. Acești oameni muncesc din greu, rezolvând pas cu pas probleme incredibil de complexe, și reușind uneori doar progrese foarte modeste. Ei sunt deștepți, sociabili, și pe baza experiențelor mele, niște oameni înțelegători și atenți.

Aceștia muncesc în proporție covârșitoare pentru nouă companii uriașe din domeniul tehnologiei- Google, Amazon, Apple, IBM, Microsoft și Facebook în Statele Unite, respectiv Baidu, Alibaba și Tencent în China -care construiesc inteligența artificială pentru a ne aduce tuturor un viitor mai bun, mai luminos. Cred cu tărie că liderii acestor companii sunt determinați de un simț profund al altruismului și de o dorință de a contribui la binele comunității: ei văd cu claritate potențialul pe care-l oferă inteligența artificială pentru a îmbunătăți sănătatea și longevitatea, a rezolva problemele climatice și a aduce bogăția în casele a mi-

lioane de oameni. Vedem deja beneficiile reale și pozitive ale muncii lor în toate domeniile și în viața de zi cu zi.

Problema e că există și forțe externe care acționează asupra celor nouă giganți tehnologici -și, prin extensie, asupra celor care lucrează în interiorul ecosistemului- obligându-i să-și desconsidere bunele intenții pentru viitorul tuturor. Există mulți vinovați.

În Statele Unite, solicitările permanente ale pieței și așteptările nerealiste în privința noilor produse și servicii zădărnicesc orice planuri pe termen lung. Ne așteptăm ca Google, Amazon, Apple, Facebook, Microsoft și IBM să anunțe noi produse și descoperiri legate de inteligența artificială la prezentările anuale, de parcă marile descoperiri ar putea fi planificate. Dacă aceste companii nu ne oferă produse mai strălucitoare decât cele de anul trecut, vorbim despre ele de parcă ar fi niște rateuri. Sau punem la îndoială existența inteligenței artificiale. Sau le punem în discuție liderii. Nu am lăsat niciodată acestor companii un răgaz de câțiva ani în care să se pună pe treabă fără să le pretindem să ne uluiască la intervale regulate. Doamne-ferește ca vreuna dintre aceste companii să se decidă să nu comunice nimic vreme de câteva luni-vom presupune că tăcerea lor ascunde un proiect care ne va lua prin surprindere.

Guvernul Statelor Unite nu are o strategie măreață în ce privește inteligența artificială sau viitorul nostru pe termen lung. Astfel, în locul unor strategii naționale coordonate menite să sporească capacitatea organizațională a guvernului, să ne construiască și întărească alianțele cu alte țări, și să ne pregătească armata pentru conflictele viitorului, Statele Unite au lăsat inteligența artificială la bunul plac al politicienilor care se perindă la guvernare. În loc

să finanțeze un program de bază de cercetare în inteligența artificială, guvernul federal a externalizat efectiv cercetarea și dezvoltarea sectorului privat și capriciilor bursei de pe Wall Street. În loc să considere că inteligența artificială este o ocazie prin care se pot crea noi slujbe și pune bazele creșterii economice, politicienii americani văd doar șomajul generalizat pe care l-ar putea provoca tehnologia. Ei dau la rândul lor vina pe giganții tehnologiei din Statele Unite, în loc să invite aceste companii să ia parte la cele mai înalte niveluri ale planificării strategice ale guvernului (în măsura în care ea există). Pionierii inteligenței artificiale nu au nicio alternativă în afara unei competiții constante pentru a avea o legătură directă cu voi, cu mine, cu școlile, spitalele, firmele și orașele noastre.

Suferim în Statele Unite de un miopism tragic. Trăim într-o paradigmă „prezentistă", preocupându-ne următorii ani mult mai mult decât oricare alt interval. Gândirea prezentistă pune mare preț pe realizările pe termen scurt ale tehnologiei, dar ne exonerează de răspunderea față de evoluția tehnologiei și rezultatele propriilor fapte. Uităm prea ușor că ceea ce facem în prezent ar putea avea consecințe grave asupra viitorului. E așadar de mirare că practic am externalizat viitorul inteligenței artificiale către șase companii listate la bursă, ale căror realizări sunt spectaculoase, dar care au interese financiare care nu se suprapun întotdeauna cu un rezultat optim pentru libertățile noastre, pentru comunitățile noastre și pentru idealurile noastre democratice?

În China însă, evoluția inteligenței artificiale este strâns legată de marile ambiții ale guvernului. China își construiește cu rapiditate bazele pentru a deveni autoritatea mondială de necontestat în domeniu. În luna iulie 2017, guvernul chinez a prezentat planul

de dezvoltare a următoarei generații de inteligență artificială pentru a obține până în anul 2030 supremația în domeniul inteligenței artificiale, urmând ca cifra de afaceri a industriei naționale să atingă cel puțin pragul de 150 de miliarde de dolari,[1] prin alocarea unei cote din fondul suveran de investiții către noi laboratoare de cercetare și companii nou înființate, precum și noi facultăți menite să pregătească următoarea generație de specialiști chinezi în inteligență artificială.[2] În luna octombrie a aceluiași an, președintele chinez Xi Jinping a explicat planurile sale pentru inteligența artificială și analiza datelor într-un discurs amănunțit în fața a mii de lideri ai partidului. Inteligența artificială, susține Xi, va ajuta China să devină una dintre cele mai avansate economii din lume. Economia Chinei este deja de 30 de ori mai mare decât era în urmă cu doar trei decenii. Deși Baidu, Tencent și Alibaba sunt giganți listați la bursă, ei se supun voinței Beijingului, aidoma oricărei mari companii din China.

Populația uriașă a Chinei, de 1,4 miliarde de locuitori, îi dă acestei țări controlul asupra celei mai mari, și poate chiar celei mai importante resurse naturale din epoca inteligenței artificiale: datele despre oameni. E nevoie de cantități considerabile de date pentru a îmbunătăți algoritmii de recunoaștere a tiparelor -motiv pentru care sistemele chinezești de recunoaștere facială precum Megvii sau SenseTime sunt atât de atrăgătoare pentru investitori. Toate datele pe care cetățenii chinezi le generează în timp ce vorbesc la telefon, fac cumpărături pe internet și postează fotografii pe rețelele de socializare sunt folosite de Baidu, Alibaba și Tencent pentru a crea sisteme performante de inteligență artificială. Un mare avantaj al Chinei este că nu are aceleași restricții de securitate

și protecție a datelor personale care ar putea limita progresul în acest domeniu în Statele Unite.

Trebuie să analizăm parcursul inteligenței artificiale în contextul mai larg al marilor planuri de viitor ale Chinei. În luna aprilie 2018, Xi a ținut un discurs important în care-și prezenta viziunea asupra Chinei ca super-putere mondială. Agenția de presă de stat Xinhua a publicat fragmente din discurs, în care Xi amintea o nouă rețea de administrare a spațiului cibernetic și un internet care ar „transmite informații pozitive, ar respecta orientarea politică corectă și ar îndruma opinia publică și valorile în direcția bună."[3] Regulile autoritare pe care China încearcă să le impună în restul lumii constituie o abatere de la libertatea de exprimare, economia de piață și împărțirea puterii, valori pe care le îndrăgim în lumea occidentală.

Inteligența artificială este doar o componentă a unor legi care urmăresc să controleze toate informațiile generate pe teritoriul Chinei și să monitorizeze datele propriilor cetățeni, precum și ale cetățenilor diferiților săi parteneri strategici. Una dintre aceste legi solicită tuturor companiilor din străinătate să stocheze datele referitoare la cetățenii chinezi în servere aflate pe teritoriul chinezesc. Din cauza acestei reglementări, agențiile de siguranță ale guvernului pot accesa după bunul plac aceste date. Un alt proiect – Cloud-ul poliției chineze- a fost creat pentru a monitoriza și urmări persoanele cu probleme psihice, disidenții declarați și o minoritate etnică musulmană- uigurii. În luna august 2018, Națiunile Unite au anunțat că există surse credibile conform cărora China deține milioane de uiguri în lagăre secrete în regiunile vestice îndepărtate ale Chinei.[4] Programul integrat de monitorizare

al Chinei folosește inteligența artificială pentru a depista eventuale abateri – detectând, de exemplu, dacă cineva își plătește facturile cu întârziere. Sistemul de credite sociale bazat pe inteligența artificială a fost creat, conform unui slogan din documentele oficiale de planificare, pentru a instaura o societate lipsită de probleme „permițând cetățenilor care respectă legea să prospere, împiedicându-i însă pe cei discreditați să facă chiar și un pas." [5] Pentru a putea promova „încrederea", cetățenii sunt evaluați în funcție de diferite criterii cum ar fi faptele de curaj (puncte în plus) sau amenzile de circulație (puncte în minus). Cei cu scoruri mai mici își găsesc mai greu o slujbă, o casă sau un loc în școală pentru copii. În unele orașe, fotografiile cetățenilor ascultători sunt afișate în public.[6] În alte orașe, de pildă în Shandong, cetățenii care trec strada prin locuri nepermise sunt afișați în public pe panouri digitale și pe Weibo, o rețea populară de socializare.[7] Dacă vi se pare neverosimil, aduceți-vă aminte că în China s-a putut decreta ca toate familiile să nu facă decât un singur copil pentru a-și reduce forțat populația.

Aceste politici și inițiative provin din anturajul apropiat președintelui Xi Jinping, format din persoane care în ultimul deceniu au acordat o atenție extraordinară repoziționării și reconstruirii Chinei ca hegemon mondial. China este mai autoritară în zilele noastre decât sub orice alt lider de după Mao Zedung, iar progresul în domeniul inteligenței artificiale și folosirea acesteia constituie elemente esențiale ale supraviețuirii regimului. Inițiativa Noului Drum al Mătăsii constituie o strategie geoeconomică deghizată sub forma unui plan vast de dezvoltare a infrastructurii de-a lungul fostului Drum al Mătăsii care unea odinioară China

cu Europa prin Orientul Mijlociu și Africa. China nu construiește doar poduri și autostrăzi- China exportă și tehnologie de supraveghere și strânge date, sporind influența Partidului Comunist Chinez în lumea, în fața ordinii liberale actuale. Interconectarea energetică globală este o altă inițiativă națională a lui Xi care urmărește crearea primei rețele globale de electricitate, sub umbrela Chinei. China dispune deja de o tehnologie revoluționară cu tensiuni foarte înalte, care-i permite să transporte curent electric din îndepărtatele regiuni vestice către Shanghai, și încheie diverse acorduri comerciale pentru a furniza energie electrică țărilor vecine.

Aceste planuri de viitor, dar și multe altele, constituie modalități inteligente de a obține o putere subtilă pe o perioadă lungă de timp. Ele constituie o mutare strategică de geniu a lui Xi, al cărui Partid Comunist Chinez a votat în luna martie 2018 să elimine limitele de mandat, permițându-i acestuia să rămână președinte pe viață. Scopul lui Xi e cât se poate de evident: să făurească o nouă ordine mondială în care China să devină liderul de facto. Cu toate acestea, într-o perioadă de expansiune a influenței diplomatice chineze, Statele Unite și-au întors spatele alianțelor globale de durată, iar președintele Trump a înălțat o cortină nouă, de bambus.

Viitorul inteligenței artificiale urmează două parcursuri de dezvoltare aflate adesea în conflict cu interesele omenirii. Demersurile chinezești în domeniul inteligenței artificiale sunt parte a unei tentative coordonate de a instaura o nouă ordine mondială sub bagheta președintelui Xi, în vreme ce în America principalii factori determinanți sunt forțele pieței și consumismul. Această

dihotomie reprezintă un punct nevralgic pentru noi toți. Rezolvarea ei reprezintă partea esențială a problemelor care ne pândesc în viitor, precum și motivul pentru care am scris această carte. Cei Nouă Titani ar putea urmări aceleași obiective nobile, comune – decodarea secretelor care ar permite construirea unor sisteme capabile să producă o gândire asemănătoare celei omenești – rezultatul final al acestei activități ar putea afecta însă iremediabil omenirea.

Eu cred că inteligența artificială este, la bază, o forță a binelui care va contribui la un trai mai bun pentru generațiile viitoare și ne va ajuta să atingem visurile despre viitor.

Sunt și o fire pragmatică, însă. Cu toții știm că până și oamenii cu intențiile cele mai bune pot să facă mult rău fără să vrea. În domeniul tehnologiei, și mai ales când vine vorba de inteligența artificială, trebuie să avem constant în vedere că aceste tehnologii pot fi folosite atât așa cum au fost proiectate, cât și abuzate în moduri neprevăzute. În prezent, dar și în viitorul pe care-l putem întrezări, această constatare este deosebit de importantă deoarece inteligența artificială afectează toate domeniile: economia globală, forța de muncă, agricultura, transporturile, sistemul bancar, monitorizarea mediului, educația, armata și siguranța națională. Acestea sunt motivele pentru care, dacă inteligența artificială își va păstra parcursul actual în Statele Unite și China, anul 2069 ar putea fi radical diferit față de cum îl anticipăm în 2019. Pe măsură ce structurile și sistemele care guvernează societatea ajung să depindă de inteligența artificială, vom constata că pentru mașini deciziile luate în locul nostru sunt complet rezonabile- doar că nouă ele nu ni se vor părea astfel.

Ca oameni, ne pierdem rapid înțelegerea asupra fenomenului, chiar în timp ce mașinile de abia se trezesc. Am depășit câteva pietre de hotar în evoluția tehnică și geopolitică a inteligenței artificiale, însă cu fiecare nou progres inteligența artificială devine din ce în ce mai opacă pentru publicul larg. Modul în care datele noastre sunt extrase și prelucrate devine mai puțin clar, în timp ce capacitatea noastră de a înțelege felul în care sistemele autonome iau decizii devine din ce în ce mai redusă. Am ajuns astfel la o prăpastie în ce privește înțelegerea felului în care inteligența artificială ne afectează viața cotidiană, iar această prăpastie se mărește exponențial pe măsură ce avansăm cu anii și deceniile în viitor. Misiunea acestei cărți este să aducă în atenția publicului dezbaterea privind inteligența artificială și să vă pregătească pentru ce urmează – și să vă fac să înțelegeți care vor fi consecințele în viața de zi cu zi pentru fiecare dintre noi, înainte de a fi prea târziu.

Omenirea se confruntă la propriu cu o criză existențială, pentru că nimeni nu răspunde la o întrebare simplă care s-a aflat la baza inteligenței artificiale încă de la început: ce se petrece cu o societate atunci când transferăm puterea unui sistem construit de un grup restrâns de oameni, care e menit să ia decizii în numele tuturor ? Ce se întâmplă când aceste decizii favorizează piața, sau un partid politic ambițios? Răspunsul se reflectă în oportunitățile viitoare pe care le avem, în felul în care ni se refuză accesul, în convențiile sociale din cadrul societăților noastre, în regulile pe baza cărora operează economiile noastre și chiar în felul în care ne raportăm la ceilalți.

Aceasta nu este o carte despre dezbaterile obișnuite din domeniu. Ea e atât un semnal de alarmă, cât și un tipar pentru un

viitor mai bun. Cartea pune la îndoială aversiunea față de planurile pe termen lung pe care o simțim în Statele Unite și subliniază lipsa de pregătire în domeniu pe care o vădesc firmele, școlile și guvernul nostru. Cartea oferă o imagine sinceră a strategiilor interconectate geopolitice, economice și diplomatice ale Chinei, care-și urmărește propria sa viziune asupra unei noi ordini mondiale. Cartea e și un îndemn către o conducere eroică în împrejurări deosebit de grele. Veți afla din paginile acestei cărți că viitorul nostru are nevoie de un erou.

Urmează un apel scris în trei părți. Din prima parte veți afla ce este inteligența artificială și rolul pe care-l joacă cei Nouă Titani în evoluția acesteia. Vom analiza îndeaproape și particularitățile celor nouă companii în America, respectiv ale Baidu, Alibaba și Tencent în China. În partea a doua veți descoperi trei scenarii amănunțite, plauzibile despre viitorii 50 de ani, pe măsură ce inteligența artificială progresează. Cele trei scenarii pe care le veți citi oscilează între o abordare optimistă, pragmatică și catastrofică, ele revelându-ne atât riscurile cât și oportunitățile pe care le incumbă trecerea de la inteligența artificială îngustă la inteligența artificială generală și apoi la superinteligența artificială. Aceste scenarii sunt intense- ele sunt rezultatul unor modele analitice, permițându-vă să aveți o perspectivă viscerală asupra evoluției potențiale a inteligenței artificiale și a schimbărilor care le vor suferi ca urmare viețile noastre. În partea a treia voi prezenta soluții tactice și strategice la toate problemele pe care le-am identificat în scenarii, dar și un plan concret de a reporni prezentul. Partea a treia este menită să ne redea imboldul de a acționa, ea conținând recomandări concrete

pentru guvernele noastre, liderii celor nouă titani, și chiar pentru voi.

Orice persoană aflată în viață în ziua de azi poate juca un rol esențial în viitorul inteligenței artificiale. Deciziile pe care le luăm acum în privința inteligenței artificiale- chiar și cele aparent banale- vor modifica pentru totdeauna parcursul istoriei omenirii. Pe măsură ce mașinile se trezesc, ne-am putea da seama că, în ciuda speranțelor și ambițiilor noastre altruiste, sistemele inteligenței artificiale ar putea reprezenta o catastrofă pentru omenire.

Nu e însă obligatoriu ca lucrurile să se petreacă astfel.

Cei Nouă Titani nu sunt personajele negative ale acestei povești. De fapt, aceste companii reprezintă cea mai bună speranță a noastră pentru viitor.

Dați pagina. Nu putem să stăm și să așteptăm viitorul. Inteligența artificială e deja printre noi.

PARTEA I

Spiritul din aparat

CAPITOLUL UNU

MINTEA ȘI MAȘINILE: O FOARTE SCURTĂ ISTORIE A INTELIGENȚEI ARTIFICIALE

Originile inteligenței artificiale moderne se întind în urmă cu sute de ani, cu mult timp înainte ca cei Nouă Titani să construiască agenți ai inteligenței artificiale cu nume ca Siri, Alexa sau echivalentul lor chinezesc, Tiān Māo. În tot acest timp, inteligența artificială nu a avut o singură definiție, așa cum au alte tehnologii. Descrierea concretă a inteligenței artificiale nu e atât de ușoară, și asta pentru că ea reprezintă multe lucruri diferite, chiar și în timp ce domeniul continuă să crească. Ceea ce se considera în anii '50 inteligență artificială- un calculator capabil să efectueze împărțiri- nu ni se mai pare vreo tehnologie avansată în zilele noastre. Acest fenomen este cunoscut sub numele de „paradoxul ciudat" de îndată ce noile tehnologii sunt inventate și devin parte a vieții cotidiene, ele devin invizibile pentru noi. Nu ni le mai reprezentăm ca fiind parte a inteligenței artificiale.

În forma sa cea mai elementară, inteligența artificială reprezintă un sistem capabil să ia decizii autonome. Activitățile desfășurate de inteligența artificială copiază sau imită inteligența omenească, sub

forma recunoașterii sunetelor, obiectelor, rezolvării problemelor, înțelegerii limbajului și a folosirii strategiei pentru atingerea unor obiective. Unele sisteme sunt gigantice, putând efectua rapid milioane de calcule, în vreme ce altele sunt dedicate unei singure misiuni, cum ar fi depistarea cuvintelor obscene din e-mailuri.

Am revenit întotdeauna la același set de întrebări: sunt mașinile în stare să gândească? Ce ar însemna ca o mașină să „gândească"? Ce înseamnă pentru noi să gândim? Ce este gândirea? Cum am putea ști- cu siguranță, și fără îndoială- că avem gânduri originale? Aceste întrebări ne-au însoțit pe parcursul secolelor și ele sunt esențiale atât în istoria cât și în viitorul inteligenței artificiale.

Problema studierii felului în care atât mașinile cât și oamenii gândesc e faptul că cuvântul „a gândi" este strâns legat de noțiunea de „minte". Dicționarul Merriam-Webster definește „a gândi" ca fiind „a forma sau a avea în minte" în vreme ce dicționarul Oxford explică că înseamnă „a folosi propria minte în mod activ pentru a înlănțui idei." Dacă căutăm cuvântul „minte", atât Merriam-Webster cât și Oxford oferă o definiție în contextul „conștienței." Dar ce este conștiența? Conform ambelor dicționare, ea reprezintă starea de a fi conștient și de a reacționa. Diferite grupuri – psihologi, neurobiologi, filosofi, teologi, eticieni și informaticieni –abordări separate când vine vorba de noțiunea de gândire.

Când folosiți Alexa pentru a găsi o masă la restaurantul preferat, atât voi cât și ea sunteți conștienți și reactivi când vine vorba de a mânca, chiar dacă Alexa nu a simțit niciodată senzația pe care o provoacă dinților textura unui măr crocant sau înțepăturile efervescente ale apei minerale pe limbă sau cum se lipește untul de arahide de cerul gurii. Întrebați-o pe Alexa să vă descrie calitățile

acestor mâncăruri, și ea vă va oferi detalii care oglindesc propriile voastre experiențe. Alexa nu are gură- și atunci cum ar putea ea să perceapă la fel ca voi mâncarea?

Sunteți o persoană unică din punct de vedere biologic, ale cărei glande salivare și papile gustative nu sunt poziționate la fel cu ale mele. Cu toate acestea, amândoi am învățat ce este un măr, caracteristicile generale ale gustului de măr, textura lui și mirosul lui. Pe parcursul vieții, am învățat să recunoaștem ce este un măr prin consolidarea informației- cineva ne-a învățat cum arată un măr, la ce servește, și prin ce se deosebește de alte fructe. După aceea, cu timpul și fără să ne dăm seama conștient de asta, sistemele noastre biologice autonome de recunoaștere a tiparelor au început să poată să stabilească cu mare precizie dacă un obiect este un măr, chiar dacă avem la dispoziție doar câteva dintre elementele necesare. Dacă vedeți un contur alb-negru, bidimensional al unui măr, veți ști ce este -chiar dacă nu are miros, gust, textură și toate celelalte date pe baza cărora creierul vostru poate identifica un măr. Felul în care voi și Alexa ați învățat despre mere e mai asemănător decât credeți.

Alexa e competentă, dar e ea *inteligentă?* E nevoie ca percepția ei mașinală să bifeze toate calitățile percepției omenești pentru ca noi să putem accepta că felul ei de „gândire" îl oglindește fidel pe al nostru? Dr. Benjamin Bloom, cercetător în psihologia educațională, și-a petrecut cea mai mare parte a carierei academice studiind și clasificând stările de gândire. El a publicat în anul 1956 taxonomia lui Bloom, care definea obiectivele de învățare și nivelurile realizărilor obținute în educație. Stratul de bază îl reprezintă reținerea datelor și conceptelor elementare, urmat apoi de înțelegerea ideilor; aplicarea cunoștințelor în situații noi; analizarea

informațiilor prin experimentare și conexiuni între idei; evaluarea, susținerea și criticarea informației; totul culminând cu crearea unor opere originale. Din copilărie, ne concentrăm mai întâi pe înmagazinare și înțelegere. Spre exemplu, trebuie mai întâi să învățăm că o sticlă conține lapte înainte de a înțelege că sticla respectivă are o parte din față și una din spate, chiar dacă nu o putem vedea.

Această ierarhie se regăsește și în felul în care calculatoarele învață. În 2017, un sistem de inteligență artificială denumit Amper a compus și produs muzică originală pentru un album intitulat I AM AI. Instrumentația și percuția au fost generate de Amper, folosind parametrii inițiali precum genul muzical, dispoziția sau lungimea melodiei, pentru a crea un cântec în doar câteva minute. Un artist viu, Taryn Southern, a colaborat cu Amper la crearea albumului, iar rezultatul a fost o baladă emoționantă denumită „Break Free" care a strâns peste 1,6 milioane de vizualizări pe YouTube și a devenit un hit la radio. Înainte ca Amper să poată crea acest cântec, a trebuit mai întâi să învețe care sunt caracteristicile unei balade, precum și date concrete, referitoare la lungimile notelor și ritmurilor și cum să recunoască mii de tipare în muzică (acorduri, accente ritmice, secvențe armonice).

Genul de creativitate demonstrată de Amper reprezintă vârful taxonomiei lui Bloom, dar avem aici de-a face doar cu un proces învățat mecanic? E un exemplu de creativitate omenească? Sau e o creativitate în cu totul alt sens? Se gândea Amper la muzică, așa cum o poate face un compozitor viu? S-ar putea susține că „creierul" lui Amper – o rețea neurală formată din algoritmi și date într-un recipient – nu e chiar atât de diferit de creierul lui Beethoven, constituit din neuroni organici care folosesc date și re-

cunosc tipare într-un recipient, adică la el în cap. A trecut Amper oare printr-un proces creativ complet diferit de cel al lui Beethoven atunci când a compus Simfonia a 5-a, cea care începe celebru prin ta-ta-ta-TAM, ta-ta-ta-TAM înainte de a trece de la o gamă majoră la una minoră? Aceste prime patru note sunt urmate de o secvență armonică, de game, arpegii și alte ingrediente brute ale oricărei compoziții. Dacă ascultați cu atenție *scherzoul,* înainte de final, veți vedea un împrumut evident din simfonia nr. 40 a lui Mozart, scrisă cu 20 de ani în urmă, în 1788. Mozart a fost influențat de rivalul lui, Antonio Salieri, și de prietenul său, Franz Joseph Haydn, care au fost la rândul lor influențați de operele unor compozitori mai vechi precum Johann Sebastian Bach, Antonio Vivaldi sau Henry Purcell, care au scris muzică de la mijlocul secolului XVII până la mijlocul secolului XVIII. În muzica lor se pot detecta pasaje din compozitori și mai vechi, de la 1400-1600, cum ar fi Jacques Arcadelt, Jean Mouton și Johannes Ockeghem. *Și ei* au fost influențați de primii compozitori medievali – și am putea continua cu influența până la prima compoziție intitulată epitaful Seikilos, gravat în secolul I e.n. într-o coloană de marmură într-un mormânt din Turcia zilelor noastre. Am putea continua chiar și mai mult în timp, până la primele flaute primitive, sculptate din os și fildeș acum 43.000 de ani. Chiar și înaintea acestei perioade, cercetătorii cred că strămoșii noștri au cântat înainte de a vorbi.[8]

Felul în care gândim este rezultatul a milioane de ani de evoluție. Felul în care funcționează inteligența artificială modernă se bazează în mod similar pe un lung parcurs al evoluției, până la primii matematicieni, filosofi și oameni de știință ai antichității. Deși ar putea părea că omenirea și mașinile au avut parcursuri diferite,

evoluția noastră a fost întotdeauna strâns legată de a lor. *Homo sapiens* a învățat să-și cunoască mediul, a transmis anumite trăsături generațiilor viitoare, s-a înmulțit și diversificat din cauză că a inventat tehnologii avansate, cum ar fi agricultura, uneltele de vânătoare sau penicilina. A fost nevoie de 11.000 de ani pentru ca de la cei 6 milioane de locuitori ai planetei din neolitic să se ajungă la populația de 7 miliarde din zilele noastre.[9] Ecosistemul în care evoluează sistemele inteligenței artificiale – învățarea, datele, algoritmii, procesoarele, aparatele și rețelele neurale – se îmbunătățește exponențial. Nu e decât o chestiune de decenii până când sistemele inteligenței artificiale se vor răspândi și integra în toate aspectele vieților noastre.

Dacă Alexa poate percepe un măr la fel ca noi, și dacă muzica lui Amper chiar e „originală" – acestea sunt întrebări care vizează mai mult felul în care definim gândirea. Inteligența artificială din ziua de azi reprezintă un amalgam la care au contribuit de-a lungul mileniilor filosofi, matematicieni, oameni de știință, roboticieni, artiști și teologi. Misiunea lor – dar și a noastră, în acest capitol – este să înțeleagă legătura dintre gândire și recipientele acesteia. Care este legătura dintre mintea omenească și mașinile construite de cei Nouă Titani în China și Statele Unite-*dacă nu cumva aceste mașini sunt construite în ciuda minții omenești*?

Există o minte în interiorul mașinii?

Primele baze ale inteligenței artificiale provin din Grecia antică, odată cu originile filosofiei, logicii și matematicii. În multe din scrierile lui Platon, Socrate spune „cunoaște-te pe tine însuți", iar el susținea că pentru a putea să deveniți mai buni și a lua de-

cizii corecte, trebuie mai întâi să vă cunoașteți caracterul. Printre alte contribuții, Aristotel a inventat logica silogistică și primul sistem formal de raționament deductiv. Cam în aceeași perioadă, matematicianul Euclid a inventat o cale pentru a găsi cel mai mare numitor comun între două numere, creând astfel primul algoritm. Din opera lor s-au născut două idei importante: că anumite sisteme fizice pot opera pe baza unui set de reguli logice, și că însăși gândirea omenească ar putea fi un sistem simbolic. De aici au pornit sute de ani de investigații ale filosofilor, teologilor și oamenilor de știință. E corpul un aparat complex? Un întreg compus din sute de alte sisteme care funcționează laolaltă, ca o pendulă? Dar mintea? E și ea o mașinărie complicată? Sau poate ceva cu totul diferit? Nu aveam cum să dovedim sau să infirmăm existența unui algoritm divin, sau o legătură între minte și lumea fizică.

În anul 1560, un ceasornicar pe nume Juanelo Turriano a creat un mic călugăr drept ofrandă adusă bisericii, în numele regelui Filip al II-lea al Spaniei, al cărui fiu își revenise în mod spectaculos după un traumatism la cap.[10] Acest călugăr avea puteri uluitoare – putea să meargă pe masă, să ridice un crucifix și un rozariu, să se bată în piept în semn de penitență și să-și miște buzele ca și cum s-ar ruga. El a fost primul *automaton* – o reprezentare mecanică a unei ființe vii. Deși cuvântul „robot" nu exista încă, călugărul reprezenta o invenție remarcabilă, care trebuie că a șocat și încurcat publicul din vremea aceea. Nimeni nu și-ar fi putut probabil închipui că un mic automat ar putea într-un viitor îndepărtat nu doar să imite mișcările cele mai elementare ci ar putea să înlocuiască oamenii în fabrici, laboratoare sau conversații de bucătărie.

Micuțul călugăr a inspirat prima generație de roboticieni, al căror obiectiv a fost să creeze mașinării din ce în ce mai complexe care să imite oamenii: automatele au devenit curând capabile să scrie, danseze sau picteze. Aici a intervenit un grup de filosofi care au început să se întrebe ce anume înseamnă a fi om. Dacă este posibil să construim automate care imită comportamentul omenesc, atunci sunt oamenii automate de fabricație divină? Sau suntem sisteme complexe capabile de rațiune și gândire proprie?

Filosoful politic englez Thomas Hobbes a descris rațiunea omenească drept o formă de calcul în opera sa *De Corpore*, parte a măreței sale trilogii despre științele naturale, psihologie și politică. Acesta scria în 1655: „Prin rațiune, înțeleg calcul. Iar a calcula înseamnă a aduna multe lucruri împreună în același timp, sau a ști ce rămâne când scădem un lucru dintr-un altul. A raționa este prin urmare același lucru cu a aduna sau a scădea."[11] Dar cum am putea ști dacă avem și liber arbitru în tot acest timp?

În timp ce Hobbes își scria prima parte a trilogiei, filosoful francez René Descartes publica opera *Meditații metafizice*, întrebându-se dacă putem ști cu siguranță că ceea ce percepem este adevărat. Cum putem să ne verificăm conștiența? De ce dovezi am avea nevoie pentru a concluziona că gândurile ne aparțin, și că lumea din jurul nostru e reală? Descartes era un raționalist, crezând că putem obține faptele prin deducere. El a pus bazele unui experiment celebru de gândire. El le-a cerut cititorilor să-și imagineze un demon care creează intenționat o iluzie a lumii lor. Dacă senzația fizică, senzorială a cititorului de a înota într-un lac nu ar fi nimic altceva decât un construct al demonului, atunci acesta nu ar putea să *știe* cu adevărat că înoată. În perspectiva lui

Descartes, însă, dacă cititorul este înzestrat cu conștientizarea propriei existențe, atunci acesta îndeplinește criteriile pentru a fi o ființă care cunoaște. „Sunt, exist, ori de câte ori rostesc asta, sau mintea mea concepe, atunci trebuie că e adevărat," scria el.[12] Cu alte cuvinte, faptul că existăm este dincolo de orice îndoială, chiar dacă în mijlocul nostru s-ar afla un demon care să vrea să ne amăgească. Sau, *gândesc, deci exist.*

Mai târziu, în opera *De Homine*, Descartes susținea că oamenii ar putea construi probabil un automat – în exemplul lui, un mic animal – care să nu poată fi distins de original. Chiar dacă am putea cândva crea un om mecanic, acesta nu ar putea fi credibil, susținea Descartes, pentru că nu ar avea minte, prin urmare nici suflet. Spre deosebire de oameni, un aparat nu ar putea îndeplini criteriul cunoașterii – nu ar putea fi niciodată conștient de sine așa cum suntem noi. Pentru Descartes, conștiința se manifestă intern – sufletul era fantoma din aparatul complicat al corpului nostru. [13]

Cu câteva decenii mai târziu, matematicianul și filosoful Gottfried Wilhelm von Leibniz a studiat noțiunea conform căreia sufletul omenesc ar fi la rândul său programat, susținând că mintea însăși reprezintă doar un recipient. Dumnezeu ar fi creat sufletul și corpul în mod armonios. Corpul ar putea fi o mașinărie complexă, însă ea este programată conform unor instrucțiuni divine. Mâinile noastre se mișcă atunci când vrem noi, dar nu noi am creat sau inventat toate mecanismele care permit această mișcare. Dacă simțim durerea sau plăcerea, aceste senzații sunt rezultatul unui sistem pre-programat, al unei linii neîntrerupte de comunicare între minte și corp.

Leibniz a venit cu propriul lui exercițiu de gândire pentru a demonstra că gândirea și percepția sunt legate inextricabil de

esența omenească. Imaginați-vă că intrați într-o moară. Clădirea este un recipient care adăpostește aparate, materii prime și muncitori. Este un sistem un sistem complex cu părți care operează armonios pentru atingerea unui singur scop, dar nu ar putea fi niciodată o minte. „Nu găsim acolo decât roți dințate și mânere care se activează unul pe celălalt, dar nimic care să poată explica percepția" scria Leibniz. „Prin urmare, percepția trebuie căutată în substanțele simple, și niciodată în lucrurile complexe cum ar fi aparatele." Argumentul lui era că indiferent cât de avansată ar fi fost moara, aparatul sau automatele, oamenii nu vor putea niciodată să construiască o mașină capabilă să gândească sau să perceapă.[14]

Cu toate acestea, Leibniz era fascinat de noțiunea reproducerii anumitor aspecte ale gândirii. Cu câteva decenii în urmă, un scriitor englez obscur pe nume Richard Braithwait, care a scris câteva cărți despre comportarea în societate a adus aminte în treacăt „calculatoarele", referindu-se la oamenii care se pricep să calculeze rapid și precis.[15] Între timp, matematicianul și inventatorul francez Blaise Pascal, care a pus bazele probabilității așa cum o cunoaștem în zilele noaste, se interesa de automatizarea calculelor. Pascal își urmărise tatăl care își calcula meticulos impozitele și voia să-i ușureze procesul. Prin urmare, Pascal a început să lucreze la un calculator automat, cu roți mecanice și cadrane mobile.[16] Calculatorul funcționa, el inspirându-l pe Leibniz să-și corecteze gândirea: mașinile nu vor avea niciodată un suflet, cu toate acestea, ar putea deveni cândva posibil să construim o mașină capabilă de gândire logică la nivelul omului. În 1673, Leibniz a descris un „Calculator bazat pe pași", un nou tip de mașină de calcul care lua decizii în baza unui sistem binar.[17] Mașina era un soi de masă de

biliard, cu mingi, găuri și canale, iar mașina deschidea găurile folosind o serie de 1 (deschis) și 0 (închis).

Calculatorul teoretic bazat pe pași al lui Leibniz a pus bazele unor noi teorii, printre care se numără și noțiunea că dacă gândirea logică ar putea fi redusă la simboluri și, ca urmare, ar putea fi analizată ca un sistem de calcul, și dacă problemele geometrice ar putea fi rezolvate folosind simboluri și numere, atunci totul ar putea fi redus la biți -inclusiv comportamentul omenesc. Aceasta reprezenta o abatere semnificativă față de filosofii din vechime: mașinăriile viitorului ar putea reproduce gândirea omenească fără să încalce providența divină. Gândirea nu presupunea în mod obligatoriu percepție, simțuri sau un suflet. Leibniz își imaginase un calculator capabil să rezolve probleme generale, chiar și cele care nu sunt de ordin matematic. Și el a presupus că limbajul se poate reduce la nivelul conceptelor elementare ale matematicii și științei, ca parte a unui limbaj universal.[18]

Urmează pur și simplu mintea și aparatul un algoritm?

Dacă Leibniz avea dreptate – iar oamenii sunt niște aparate cu suflete care vor inventa cândva mașini lipsite de suflet capabile de o gândire sofisticată, inimaginabilă – atunci ar putea exista o clasă binară de aparate pe pământ: noi și ele. Dezbaterea era însă de abia la început.

În 1738, Jacques de Vaucanson, artist și inventator, a construit pentru Academia Franceză de Științe o serie de automate, printre care se număra și o rață complexă, verosimilă. Ea nu doar că imita

mişcările unei raţe vii, dând din aripi şi mâncând boabe – ci putea imita chiar şi digestia. Filosofii au avut ce diseca: dacă arăta ca o raţă, măcănea ca o raţă, era oare o raţă? Dacă credem că raţa are un altfel de suflet, ar fi suficient pentru a putea dovedi că raţa avea conştiinţă de sine, cu tot ceea ce ar presupune asta?

Filosoful scoţian David Hume a respins ideea că recunoaşterea existenţei constituie în sine dovada conştienţei. Spre deosebire de Descartes, Hume era un empirist. El a dezvoltat un nou cadru de cercetare bazat pe fapte observabile şi argumente logice. În timp ce de Vaucanson îşi prezenta raţa care digeră – şi cu mult timp ca vreun om să fi adus în discuţie inteligenţa artificială – Hume scria, în *Un tratat despre natura umană*:“ Raţiunea este, şi ar trebui să fie, roaba pasiunilor.“ În acest caz, Hume înţelegea prin „pasiuni“ „motivaţii iraţionale“ şi faptul că stimulii se află la baza comportamentului nostru, nu logica abstractă. Dacă impresiile sunt doar felul în care percepem lucrurile pe care le putem vedea, atinge, simţi, gusta şi mirosi, iar ideile sunt percepţii ale lucrurilor cu care nu intrăm direct în contact, Hume era de părere că existenţa şi înţelegerea lumii din jurul nostru se bazează pe un construct al percepţiei umane.

Odată cu noile progrese ale automatelor, care deveneau din ce în ce mai realiste, şi pe măsură ce ideea calculatoarelor ca maşini de gândit era luată din ce în ce mai serios, doctorul şi filosoful francez Julien Offray de La Mettrie a efectuat un studiu radical – şi scandalos – al oamenilor, animalelor şi automatelor. Într-un articol publicat anonim în 1747, acesta concluziona că oamenii sunt foarte asemănători cu animalele, şi că dacă un hominin ar fi „pregătit cum trebuie“, acesta ar putea stăpâni limbajul oame-

nilor. La Mettrie e de părere și că oamenii și animalele sunt doar niște aparate, mânate de instinct și experiență. „Corpul omenesc e o mașină care-și întoarce singură arcurile; ...sufletul e doar un principiu al mișcării, sau o parte sensibilă și materială a creierului.“ [19]

Ideea că oamenii ar fi doar niște aparate bazate pe materie – niște rotițe care îndeplinesc un set de funcții – presupune că nu suntem unici, sau speciali. Ea mai presupune și faptul că am putea fi programabili. Dacă ar fi adevărat, și dacă până în acest moment am putut crea rațe și călugări care să arate ca originalul, atunci ar trebuie să rezulte că într-o bună zi, oamenii s-ar putea reproduce pe ei înșiși – și să construiască diverse aparate inteligente, capabile de gândire.

Putem construi o mașină care să gândească?

În anii 1830, matematicienii, inginerii și oamenii de știință au început să meșterească, sperând că vor putea crea mașini capabile să facă aceleași calcule ca și „calculatoarele“ umane. Matematicienul Ada Lovelace și omul de știință Charles Babbage au inventat o mașină denumită „Motorul diferențial“ și apoi au imaginat un „Motor analitic“ care folosea o serie de etape prestabilite pentru a rezolva probleme matematice. Lui Babbage nu i se părea că aparatul ar putea fi capabil de altceva decât de a calcula numerele. Lovelace a fost cea care, în notele de subsol ale unui articol științific pe care îl traducea, s-a abătut în mod genial speculând că o variantă mai puternică a Motorului ar putea fi folosită în alte

moduri.[20] Dacă maşina ar putea manipula simboluri, iar acestea ar putea la rândul lor să reprezinte diferite lucruri (spre exemplu note muzicale) atunci Motorul ar putea fi folosit pentru a „gândi" în afara domeniului matematicii. Deşi ea nu credea că un calculator ar putea vreodată să devină capabil să gândească pe cont propriu, ea şi-a imaginat un sistem complex care putea urma nişte instrucţiuni şi să imite o bună parte din activităţile zilnice ale oamenilor. Unora li se părea complet banal la vremea respectivă, însă Ada reuşise să scrie primul program complet pentru un calculator puternic al viitorului- cu decenii înainte ca măcar becul să fi fost inventat.

O sută de mile mai la nord de universitatea Cambridge unde lucraseră Lovelace şi Babbage, un tânăr matematician auto-didact pe nume George Boole se plimba pe un câmp în Doncaster şi a avut o revelaţie spontană, hotărându-se să-şi dedice viaţa explicării logicii gândirii omeneşti.[21] În urma acelei plimbări a rezultat algebra booleană, care constituie o manieră de a simplifica expresiile logice folosind simboluri şi numere (de ex., „şi", „sau" şi „nu"). Astfel, de exemplu, dacă calculăm „adevărat *şi* adevărat", rezultatul ar fi adevărat, ceea ce ar corespunde comutatoarelor unui calculator. Boole a avut nevoie de două decenii pentru a-şi pune pe hârtie teoriile. Şi a mai fost nevoie de încă o sută de ani pentru ca cineva să-şi dea seama că logica şi probabilitatea booleană ar putea ajuta calculatoarele să treacă de la automatizarea unor calcule aritmetice către a deveni maşini de gândire mai complexe. Nu exista o cale de a construi o maşină gânditoare – procesele, materialele şi energia nu erau încă disponibile-aşadar teoria nu a putut fi verificată.

Saltul de la mașinile de gândire teoretică la computere care să înceapă să imite gândirea omenească s-a produs în anii 1930, odată cu apariția a două articole revoluționare: „A Symbolic Analysis of Switching and Relay Circuits" de Claude Shannon respectiv „On Computable Numbers, With an Application to the *Entscheidungsproblem*" de Alan Turing. Ca student în inginerie electrică la MIT, Shannon a făcut un curs opțional de filosofie- o abatere neobișnuită. *Legile gândirii* ale lui Boole au devenit principala sursă pentru teza lui Shannon. Îndrumătorul lui, Vannevar Bush, l-a încurajat să aplice logica booleană circuitelor fizice. Bush construise o variantă avansată a Motorului Analitic al lui Lovelace și Babbage – prototipul lui se numea Analizorul Diferențial – iar structura era întrucâtva improvizată. La momentul respectiv nu exista o teorie sistematică prin care să se impună cum ar trebui legate circuitele electrice. Descoperirea lui Shannon a constat în transpunerea logicii simbolice a lui Boole asupra circuitelor electrice, apoi în explicarea felului în care logica booleană ar putea fi folosită pentru crearea unui circuit în care să se adauge 1 și 0. Shannon și-a dat seama că pentru un calculator e nevoie de două straturi: cel fizic (recipientul) și cel logic (codul).

În timp ce Shannon încerca să aplice logica booleană unor circuite fizice, Turing testa mașina universală de traducere a lui Leibniz care ar putea reprezenta toată cunoașterea matematică științifică. Turing urmărea să dovedească ceea ce la vremea respectivă se numea *Entscheidungsproblem*, sau „problema alegerii." Pe scurt, problema poate fi rezumată astfel: nu există niciun algoritm care să stabilească dacă o afirmație matematică arbitrară este adevărată sau falsă. Răspunsul era negativ. Turing a putut dovedi

că nu există niciun algoritm, însă, ca produs secundar, el a descoperit un model matematic al unei mașini de calcul care ar putea fi folosită în orice domeniu.[22]

Și de aici s-a schimbat totul. Turing și-a dat seama că un program și datele acestuia ar putea fi stocate într-un calculator – din nou, în anii 30 această idee era una radicală. Până în acel moment, toată lumea era de acord că mașina, programul și datele erau independente. Mașina universală a lui Turing a reușit să explice pentru prima dată de ce cele trei sunt interconectate. Din punct de vedere mecanic, logica aflată la baza circuitelor și comutatoarelor putea de asemenea să fie stocată sub forma programului sau a datelor. Gândiți-vă la importanța acestor teorii. Recipientul, programul și datele constituiau o singură entitate, asemănătoare oamenilor. Și noi suntem recipiente (corpurile noastre), programe (funcții celulare autonome) și date (ADN-ul nostru, combinat cu informațiile senzoriale directe și indirecte).

Între timp, lunga tradiție a automatelor care începuse cu 400 de ani în urmă cu un mic călugăr care se ruga s-a intersectat în final cu Shannon și Turing. Compania americană Westinghouse a construit un robot denumit „Elektro Moto-Omul Westinghouse" pentru Expoziția Universală din 1939. Era un gigant rudimentar, colorat auriu, și cu roți la picioare. Avea 48 de circuite care funcționau pe platforma unei centrale telefonice. Elektro răspundea, pe baza unor mesaje preînregistrate pe un patefon, la comenzile vocale date printr-un receptor de telefon. Era un computer antropomorf, capabil de decizii rudimentare -spre exemplu ce să spună- fără o implicare directă, în timp real a unui om.

Judecând după titlurile ziarelor, nuvelele SF și buletinele de știri din acea perioadă, e clar că oamenii au fost luați prin surprin-

dere, șocați și preocupați de această evoluție. Pentru ei, părea că începuse peste noapte epoca „mașinilor gânditoare". Scriitorul SF Isaac Asimov a publicat „Liar!", o povestire pătrunzătoare în numărul din mai 1941 al revistei *Astounding Science Fiction*. Era o reacție la studiile avangardiste de care devenise conștient, formulând în cadrul acesteia cele Trei Legi ale Roboticii:

1. Un robot nu poate răni un om, sau prin inacțiune să permită ca un om să sufere.
2. Un robot trebuie să asculte ordinele pe care i le dau oamenii, mai puțin în cazurile în care aceste ordine s-ar afla în conflict cu Prima Lege.
3. Un robot trebuie să-și protejeze existența atât timp cât această protecție nu intră în conflict cu Prima sau a Doua Lege.

Ulterior, Asimov a adăugat o lege „Zero", care să le supună pe toate celelalte: „Un robot nu poate face rău omenirii, sau prin inacțiune, să permită ca omenirea să sufere."

Dar ce ar gândi o mașină gânditoare?

În 1943, cercetătorii în psihiatrie Warren McCulloch și Walter Pitts de la University of Chicago au publicat un articol important „Calculul logic al ideilor imanente activității nervoase" prin care descriau un nou tip de sistem care modelează neuronii biologici într-o arhitectură neurală de rețea pentru inteligență. Dacă recipientele, programele și datele sunt interconectate, așa cum susținea Turing, și dacă și oamenii ar fi niște recipiente la fel de elegant

create și capabile de procesarea datelor, atunci rezultă că e posibilă construirea unei mașini de gândit dacă aceasta ar fi modelată folosind organul responsabil pentru gândire- creierele noastre. Ei au propus o teorie a minții și creierului bazată pe calcul, o „rețea neurală." În loc să ne concentrăm asupra mașinii ca hardware și a programului ca software, ei și-au imaginat un nou tip de sistem simbiotic capabil să înghită cantități mare de date, așa cum facem și noi oamenii. Computerele nu erau încă suficient de puternice la acea vreme pentru a putea testa această teorie, însă articolul lor i-a inspirat pe alții să încerce să construiască un nou sistem inteligent de calcul.

Legătura dintre sistemele inteligente de calcul și procesul decizional autonom a devenit mai clară odată ce savantul John von Neumann, specialist în informatică, fizică și matematică a publicat un tratat impresionant de matematici aplicate. Scrisă împreună cu economistul de la Princeton Oskar Morgenstern în 1944, carte de 641 de pagini explica cu lux de amănunte felul în care teoria jocului stătea la baza tuturor deciziilor economice. Această lucrare a dus la colaborarea dintre von Neumann și armata americană, care dezvolta la acel moment un nou tip de calculator electric denumit Electronic Numerical Integrator and Computer, sau ENIAC, pe scurt.[a] La început, instrucțiunile ENIAC erau inserate fizic în sistem, ceea ce însemna că fiecare reprogramare presupunea recrearea fizică a conexiunilor. Inspirat de Turing, McCulloch și Pitts, von Neumann a elaborat o metodă de a stoca programele în calculator. Aceasta a reprezentat trecerea de la prima eră a calculatoarelor (tabularea) către o nouă epocă a sistemelor programabile.

---

a Integrator Numeric Electronic și Calculator (n.t.)

Turing însuși lucra la o rețea neurală alcătuită din calculatoare cu programe stocate pe sistemele respective. În 1949, publicația *The London Times* relata afirmația lui Turing: „Nu văd de ce mașina nu ar putea concura în vreunul dintre domeniile care sunt în prezent apanajul intelectului omenesc și să nu ajungă eventual să concureze de pe poziții de egalitate. Nu cred că puteți trage linie când vine vorba de sonete, deși poate că această comparație e ușor nedreaptă pentru că un sonet scris de o mașină va fi mai bine apreciat de o altă mașină." După un an, într-un articol publicat în jurnalul de filosofie *Mind*, Turing vorbea despre întrebările ridicate de Hobbes, Descartes, Hume și Leibniz. În acest articol, el a propus o teză și un test: dacă într-o bună zi, un calculator va deveni capabil să răspundă la întrebări de o manieră care să nu poată fi distinsă de cea omenească, atunci acel calculator va „gândi." Probabil că ați auzit de acest articol sub o altă denumire: testul Turing.

Articolul începe cu o întrebare devenită celebră, o întrebare pe care și-au pus-o și la care au răspuns atâția filosofi, teologi, matematicieni și oameni de știință înaintea lui: „Poate o mașină să gândească?" Turing însă, conștient de dezbatere veche de secole despre minte și mașini a respins întrebarea, susținând că e prea vastă pentru a genera o discuție fertilă. „Mașină" și „a gândi" sunt termeni ambigui, care lasă prea mult loc interpretărilor subiective. (La urma urmei, existau deja patru secole în care se scriseseră cărți și articole despre înțelesul acestor cuvinte.)

Jocul presupunea o amăgire, iar calculatorul „câștiga" dacă putea trece drept un om. Testul se desfășoară în felul următor: există un om, o mașină, și o persoană care pune întrebări într-o cameră separată. Obiectivul jocului este ca interogatorul să-și dea

seama care dintre răspunsuri provin de la o persoană, și care provin de la mașină. La începutul jocului, interogatorul primește etichete, X și Y -neștiind însă care este cea care corespunde calculatorului- și i se permite să pună doar întrebări de genul „Poate X să-mi spună dacă X joacă șah?" La finalul acestui joc, interogatorul trebuie să-și dea seama cine era X și cine era Y. Cealaltă persoană are misiunea de a-l ajuta pe interogator să identifice mașina, iar mașina trebuie să păcălească interogatorul să creadă că e de fapt om. În privința jocului, Turing a scris:

Cred că peste vreo 50 de ani va deveni posibil să programăm niște calculatoare cu o capacitate de stocare de 10^9, să le punem să joace jocul imitației atât de bine încât un interogator mediu nu ar avea mai mult de 70% șanse să identifice corect personajele după cinci minute de întrebări.[23]

Turing era însă un om de știință și el știa că această teorie nu putea fi dovedită, cel puțin nu în cursul vieții lui. Problema s-a dovedit a fi însă nu faptul că Turing nu dispunea de dovezi empirice care să dovedească că mașinile ar putea într-o zi să gândească, și nici momentul -Turing a spus că ar fi nevoie să așteptăm până la finalul secolului XX pentru a putea desfășura acest test. „Sperăm ca mașinile să poată concura în cele din urmă cu oamenii în toate domeniile intelectuale," scria Turing.

Adevărata problemă era să facem în așa fel încât să ajungem credem că mașinile ar putea într-o bună zi să vadă, să raționeze și să își aducă aminte – și că oamenii ar putea încerca să împiedice acest progres. Astfel, colegii lui ar fi trebuit să observe cogniția lipsiți de spiritualism și să creadă în plauzibilul unor mașini inte-

ligente care, spre deosebire de oameni, să ia decizii într-un mod inconștient.

Vara și iarna inteligenței artificiale

Profesorii Marvin Minsky (matematică și neurologie) și John McCarthy (matematică) împreună cu Claude Shannon (matematician și criptograf la Bell Labs) și Nathaniel Rochester (informatician la IBM) au propus în 1955 un atelier de două luni dedicat operei lui Turing și posibilității ca mașinile să poată învăța. Teoria lor era următoarea: dacă era posibil să descriem fiecare caracteristică a inteligenței umane, atunci o mașină ar putea fi învățată să o simuleze.[24] Era însă nevoie de un grup mare, divers de experți din multe domenii diferite. Ei considerau că puteau înregistra progrese semnificative strângând laolaltă un grup interdisciplinar de cercetători și lucrând intensiv, fără întreruperi, pe durata verii.

Selecția grupului avea o importanță critică. El avea să devină rețeaua de ingineri, sociologi, informaticieni, psihologi, matematicieni, fizicieni și specialiști în cogniție care urmau să pună întrebările fundamentale și să le răspundă, întrebări despre ce înseamnă să „gândim", cum funcționează „mințile" noastre și cum putem face mașinile să învețe la fel cum o facem noi oamenii. Se intenționa ca această rețea diversă să continue să colaboreze în domeniul cercetării și să ducă acest nou domeniu mai departe. Pentru că era vorba de o abordare nouă, interdisciplinară cu obiectivul de a construi mașini care să gândească, aceștia aveau nevoie de un

nou nume prin care să-şi descrie activitatea. Au găsit o descriere ambiguă dar elegantă: inteligenţa artificială.

McCarthy a creat o listă preliminară cu 47 de experţi de care credea că avea nevoie pentru a pune bazele grupului şi ale activităţii de cercetare. A fost un proces tensionat prin care s-au stabilit toate persoanele importante care trebuiau să fie prezente la demersul teoretic şi practic al inteligenţei artificiale. Minsky era în mod special preocupat de faptul că la întâlnire aveau să lipsească două voci esenţiale-Turing, care decedase cu doi ani în urmă, şi von Neumann, suferind de cancer în fază terminală.[25]

Cu toate eforturile de a aduna un grup divers cu cea mai bună combinaţie de aptitudini complementare, ei aveau şi un minus. Toţi erau albi, deşi existau mulţi oameni de culoare care lucrau în domeniile pe care McCarthy şi Minsky voiau să le reunească Toţi erau angajaţi ai marilor companii din domeniul tehnologiei de la vremea respectivă (IBM, Bell Labs) sau proveneau de la o mână de facultăţi. Chiar dacă erau numeroase femei de geniu care contribuiseră semnificativ în domeniul ingineriei, informaticii, matematicii sau fizicii, ele au fost excluse.[26] Toţi participanţii erau bărbaţi, în afară de Gloria, soţia lui Marvin Minsky. Fără să-şi dea seama de propriile prejudecăţi, aceşti oameni de ştiinţă- care sperau să înţeleagă cum funcţionează mintea omenească, cum gândim şi cum ar putea maşinile să înveţe de la toţi oamenii-şi-au limitat drastic datele la cei care arată şi vorbesc la fel ca ei.

În anul următor, grupul s-a adunat la ultimul etaj al facultăţii de matematică din Dartmouth pentru a studia teoria complexităţii, simularea limbajului natural, reţelele neurale, legătura dintre aleatoriu şi creativitate şi maşinile care pot învăţa. În timpul săptămânii se în-

tâlneau în sala de curs pentru o discuție generală înainte de a se dedica unor sarcini concrete. Profesorii Allen Newell, Herbert Simon și Cliff Shaw au descoperit o metodă de a descoperi dovezi ale teoremelor logice și simulau procesul – un program pe care l-au denumit „Teoreticianul Logic" – într-una din sesiunile generale. Acesta a fost primul program care a mimat aptitudinile omenești de rezolvare a problemelor. (Programul a dovedit într-un final 38 din cele 52 de teoreme din *Principia Mathematica* a lui Alfred North Whitehead și Bertrand Russell, o carte de referință despre bazele matematicii.) Claude Shannon, care sugerase cu câțiva ani în urmă că ar trebui să învățăm calculatoarele să joace șah împotriva oamenilor, a avut ocazia să prezinte un prototip al acestui program, aflat încă în faza de dezvoltare.[27]

Așteptările lui McCarthy și Minsky de a progresa spectaculos în domeniul inteligenței artificiale nu s-au materializat în vara aceea la Dartmouth. Nu era suficient timp – ca să nu mai vorbim de puterea de calcul – pentru ca inteligența artificială să treacă de la teorie la practică.[28] Cu toate acestea, în vara aceea au fost stabilite trei practici care au stat la baza inteligenței artificiale așa cum avea ea să se manifeste ulterior:

1. Inteligența artificială avea să fie concepută, construită, testată și perfecționată de o colaborare între marile companii din domeniul informaticii și cercetătorii academici.
2. Progresul inteligenței artificiale este o activitate costisitoare, prin urmare fiind necesară comercializarea activității – fie colaborând prin parteneriate cu agențiile guvernamentale sau cu armata, sau prin construirea unor produse și sisteme care să poată fi vândute.

3. Cercetarea și construirea inteligenței artificiale se bazează pe o rețea interdisciplinară de cercetători, ceea ce presupune înființarea de la zero a unui nou domeniu academic. Acest lucru a însemnat și că persoanele din acest domeniu au avut tendința de a recruta oameni pe care-i cunoșteau deja, păstrând omogenitatea rețelei dar limitându-i perspectiva asupra lumii.

Tot în acea vară a mai avut loc un eveniment interesant. În timp ce grupul dezbătea întrebarea lui Turing – *Poate o mașină să gândească?* aceștia nu au putut cădea de acord asupra modalității optime de a dovedi răspunsul, respectiv de a construi o mașină care să poată învăța. Unii membri preferau o abordare biologică. Ei credeau că se pot folosi rețele neurale care să poată imprima inteligenței artificiale un bun simț și un raționament logic – că ar fi posibil ca mașinile să fie în general inteligente. Alții susțineau că ar fi imposibil ca să se creeze vreodată o copie atât de fidelă a structurilor omenești de gândire. Ei preferau în schimb o abordare inginerească. În loc să scrie comenzi pentru a rezolva probleme, un program ar putea ajuta sistemul să „învețe" dintr-un set de date. Sistemul ar face predicții pe baza datelor, iar un supervizor uman ar verifica acele răspunsuri – perfecționând mașina pe parcurs. Privită din acest unghi, capacitatea mașinilor de a „învăța" primea o definiție îngustă, care se rezuma la învățarea unei sarcini concrete, cum ar fi a juca dame.

Psihologul Frank Rosenblatt, prezent la atelierul de la Dartmouth, voia să modeleze felul în care creierul uman procesează informațiile vizuale și, prin urmare, cum învață să recu-

noască obiecte. Pe baza studiilor din acea vară, Rosenblatt a creat un sistem denumit Perceptron. El intenționa să construiască un simplu program cadru care să reacționeze la feedback. Aceasta a fost prima rețea neurală artificială care opera creând conexiuni între diferite elemente de procesare într-o structură stratificată. Fiecare neuron mecanic primea multe semnale diferite, apoi folosea un sistem de ponderare matematică pentru a stabili ce semnal să genereze. În această structură paralelă puteau fi accesați mai multe procesoare simultan- lucru care însemna nu doar că era rapidă, ci că putea procesa continuu mai multe date.

Iată care este importanța acestei structuri: deși ea nu însemna neapărat că un calculator ar putea „gândi", ea ne demonstra însă cum putem face un calculator să învețe. Noi oamenii învățăm încercând și dând greș. Pentru a cânta gama do major la pian trebuie să apăsăm pe clape în ordinea corectă. La început, degetele, urechile și ochii noștri nu au memorat tiparul corect însă dacă exersăm – dacă repetăm gama constant, corectându-ne de fiecare dată- în cele din urmă vom reuși. Când am luat lecții de pian și am cântat prost gamele, profesorul mă corecta, dar dacă le nimeream, primeam un abțibild. Abțibildul mă ajuta să înțeleg că luasem deciziile corecte atunci când cântam. Și rețeaua neurală a lui Rosenblatt funcționează la fel. Sistemul învăța cum să optimizeze răspunsul repetând de mii de ori aceleași funcții și reținea ce a învățat aplicând această cunoaștere pentru noi probleme. El antrena sistemul folosind o tehnică denumită retro-propagare. În faza inițială, un om evaluează dacă rețeaua neurală artificială lua decizia corectă. Dacă lua decizia corectă, procesul era consolidat. Dacă nu, se ajustau parametrii și se aplica un alt test.

În anii de după atelier s-au înregistrat progrese semnificative în domeniul unor probleme complexe pentru oameni, cum ar fi folosirea inteligenței artificiale pentru rezolvarea teoremelor matematice. Cu toate acestea, o sarcină simplă – cum ar fi recunoașterea limbajului – rămânea o provocare uriașă, fără soluții imediate. Înainte de cercetarea în domeniul inteligenței artificiale, mintea fusese considerată dintotdeauna o cutie neagră. Datele intrau, și din cutie ieșea un răspuns fără să existe o modalitate de a observa acest proces. Primii filosofi, matematicieni și oameni de știință spuneau că acesta era rezultatul unei creații divine. Oamenii de știință din epoca modernă știau că era rezultatul a sute de mii de ani de evoluție. Abia în anii 50, după vara de la Dartmouth, cercetătorii au început să creadă că pot decoda (cel puțin pe hârtie) cutia neagră și să observe procesele cognitive, apoi să învețe calculatoarele să imite comportamentul nostru de răspuns la stimuli.

Până în acest punct, calculatoarele fuseseră doar un instrument de automatizare a calculelor. Prima eră a informaticii, marcată de mașini care puteau să calculeze numere, lăsa loc celei de-a doua ere, cea a computerelor programabile. Acestea erau sisteme mai rapide și mai compacte care dispuneau de suficientă memorie pentru a stoca seturi de instrucțiuni. Programele puteau acum fi stocate local, și, mai ales, puteau fi scrise în limba engleză, și nu sub forma unui cod complicat pentru mașini. Devenea clar că nu aveam nevoie de automate sau de recipiente antropomorfe pentru ca aplicațiile inteligenței artificiale să devină utile. Inteligența artificială putea să se găsească într-o cutie simplă, fără trăsături omenești, și cu toate acestea să fie extrem de utilă.

Atelierul de la Dartmouth l-a inspirat pe matematicianul britanic I.J. Good să scrie despre o „mașină a hiperinteligenței" care ar putea crea mașini mai bune decât noi. Prin urmare ar rezulta o „explozie a inteligenței, iar inteligența omenească ar fi lăsată mult în urmă. Așadar, această primă mașină ultra inteligentă ar fi ultima invenție pe care ar trebui să o mai facă omul."[29]

În toată această evoluție și-a făcut intrarea și o femeie, cel puțin cu numele. Informaticianul de la MIT Joseph Weizenbaum a conceput un sistem denumit ELIZA, un program de conversație denumit după personajul din piesa *Pygmalion* de George Bernard Shaw.[30] Această evoluție a fost importantă pentru rețelele neurale și inteligența artificială pentru că a fost una din primele încercări de a procesa limbajul natural, iar programul accesa diferite scenarii create în prealabil pentru a conversa cu oameni în carne și oase. Cel mai faimos dintre acestea se numea DOCTOR[31] și imita un psiholog empatic folosind recunoașterea tiparelor pentru a oferi răspunsuri frapant de umane.

Atelierul de la Dartmouth atrăsese acum atenția comunității internaționale, ca și cercetătorii care s-au trezit pe neașteptate în lumina reflectoarelor. Erau niște staruri rock educate, care le ofereau oamenilor obișnuiți o perspectivă fantastică, nouă asupra viitorului. Vă aduceți aminte de Rosenblatt, psihologul care crease prima rețea neurală? El a declarat pentru *Chicago Tribune* că în curând mașinile nu vor avea doar programe de tip ELIZA, capabile să ofere câteva sute de răspunsuri – ci vor putea să participe la ședințe și să transcrie „ca orice secretară." El a promis nu doar că va crea cel mai mare „aparat de gândit" din lume, ci și că îl va produce în timp de câteva luni. [32]

Și ce susțineau Simon și Newell, autorii Teoreticianului Logic? Ei au început să emită previziuni îndrăznețe, radicale despre inteligența artificială, spunând că în zece ani – *adică până în 1967* – calculatoarele vor:

> bate toți campionii de șah și vor deveni campioane mondiale
> descoperi și dovedi o nouă teoremă matematică importantă
> scrie o muzică pe care și cei mai aprigi critici să o aprecieze[33]

Între timp, Minsky făcea previziuni despre o mașină inteligentă care să poată face mult mai multe decât să scrie după dictare, să joace șah sau să scrie muzică. El considera că până la finalul vieții sale, mașinile vor dobândi inteligența artificială generală – cu alte cuvinte, că vor deveni capabile de gândire complexă, se vor putea exprima prin limbaj și vor putea face alegeri.[34]

Cercetătorii de la atelierul din Dartmouth au scris articole și cărți. Au dat interviuri la radio, televiziune, în ziare și reviste. Știința era însă dificil de explicat așa că în mod frecvent explicațiile au fost distorsionate și cuvintele lor au fost scoase din context. Lăsând la o parte previziunile exagerate, așteptările publicului în privința inteligenței artificiale au devenit din ce în ce mai fantastice, și asta și pentru că povestea a fost relatată greșit. Spre exemplu, Minsky ar fi spus conform revistei *Life*: „În trei-opt ani vom avea o mașină de nivelul general de inteligență al unui om mediu. Vreau să spun că vom avea o mașină care va putea să citească din Shakespeare, să schimbe uleiul la mașină, să bârfească la muncă, să spună o glumă sau să se bată."[35] În același articol, jurnalistul se referă la Alan Turing denumindu-l „Ronald Turing." Minsky, deși entuziast, nu voia să spună că mai era doar un pas până la roboți

vorbitori ci mai degrabă oferise un răspuns arogant. În absența unui context și a explicațiilor, percepția publică asupra inteligenței artificiale a început să se distorsioneze.

Faptul că în 1968 Arthur Clarke și Stanley Kubrick au decis să facă un film despre un viitor în care mașinile vor avea nivelul general de inteligență al omului mediu nu a ajutat. Ei voiau să spună o poveste despre oameni și mașinile de gândit – iar Minsky a fost cooptat în calitate de consultant. Dacă încă nu ați ghicit, e vorba de un film pe care îl cunoașteți deja – 2001: Odiseea spațială, axat în jurul unei inteligențe artificiale pe nume HAL 9000 care a învățat simțul umorului și creativitatea de la cei care l-au creat- și amenință să ucidă pe oricine ar vrea să-l deconecteze. Unul dintre personaje, Victor Kaminski, a fost denumit în cinstea lui Minsky.

Putem spune că la mijlocul anilor 60, inteligența artificială devenise parte a spiritului vremurilor, și toată lumea fetișiza viitorul. Crescuseră și așteptările legate de succesul comercial al inteligenței artificiale datorită unui articol apărut într-o publicație obscură a industriei radio. Cu un titlu simplist ca „Cum să îndesăm mai multe componente în circuitele integrate", articolul, scris de co-fondatorul intel, Gordon Moore, postula o teorie conform căreia numărul de tranzistori care pot încăpea pe o placă de circuite la același preț se va dubla la fiecare 18-24 de luni. Această idee îndrăzneață a devenit cunoscută drept Legea lui Moore, iar la început, această teză părea să fie corectă. Computerele deveneau din ce în ce mai puternice, și capabile de o multitudine de exerciții, nu doar probleme de matematică. Comunitatea inteligenței artificiale urma să primească un sprijin, pentru că asta presupunea că teoriile lor urmau curând să poată fi verificate. Legea lui Moore

mai introducea şi posibilitatea extraordinară ca procesoarele fabricate de oameni să depăşească într-o bună zi puterile minţii omeneşti, care are o capacitate de înmagazinare limitată biologic.

Toate aceste exagerări şi acest articol au canalizat multe fonduri în direcţia inteligenţei artificiale -chiar dacă persoanele care nu făceau parte din reţeaua de la Dartmouth nu înţelegeau exact ce este inteligenţa artificială. Nu existau încă produse şi nici metode practice de a aplica reţelele neurale la scară mare sau tehnologia necesară. Fiindcă oamenii credeau acum în *posibilitatea* maşinilor gânditoare, a fost suficient pentru a obţine fonduri substanţiale din mediul privat dar şi de la guvern.

Spre exemplu, guvernul Statelor Unite a finanţat un program ambiţios în care inteligenţa artificială trebuia folosită pentru a traduce limbajul. Era în toiul Războiului Rece, iar guvernul voia să aibă la dispoziţie un sistem de traducere instant a limbii ruse care să ofere o eficienţă mai mare, reduceri de costuri şi sporirea preciziei. Se părea că programarea ar fi putut oferi soluţia. Dintr-o colaborare între Institutul de Literatură şi Lingvistică de la Universitatea Georgetown şi IBM a rezultat un prototip de aparat de traducere din rusă în engleză care avea un vocabular limitat la 250 de cuvinte şi era specializat doar în chimie organică. Demonstraţia cu public a provocat un entuziasm iar traducerea automatizată a ajuns pe prima pagina a ziarului *New York Times*- dar şi a altor 6 ziare.

Fondurile erau abundente -graţie finanţării agenţiilor guvernamentale, universităţilor şi marilor companii din domeniul tehnologiei – şi se părea pentru o vreme că nimeni nu ţinea socoteala. Cu toate acestea, dincolo de articole şi prototipuri, inteligenţa artificială nu reuşea să ofere ceea ce se prevăzuse şi promisese. S-a

dovedit că progresele semnificative reprezentau o provocare mult mai mare decât anticipaseră pionierii modernității.

La scurt timp după aceea, s-a cerut o analiză a felului în care inteligența artificială ar putea fi folosită în viața de zi și a implementării ei concrete. National Academy of Sciences a înființat o comisie consultativă la solicitarea National Science Foundation, a Departamentului Apărării și CIA. Comisia a constatat că părerile erau contradictorii în ce privește posibilitatea unui program bazat pe inteligență artificială care să efectueze traduceri în timp real, concluzionând că „nu există traduceri efectuate de aparate ale unor texte științifice generale, și nici perspective imediate pentru ca ele să apară în viitor."[36] Un raport produs ulterior pentru British Science Research Council susținea că cercetătorii au exagerat avansul tehnologic în domeniul inteligenței artificiale, oferind o prognoză pesimistă pentru toate ariile de bază de cercetare din acest domeniu. James Lighthill, un matematician de la Cambridge, era principalul autor al raportului; reproșul cel mai puternic era că aceste tehnici de debut ale inteligenței artificiale – spre exemplu, a învăța un calculator să joace dame – nu puteau să producă soluții pentru problemele mai complexe ale lumii reale.[37]

În urma acestor rapoarte, politicienii din Statele Unite și Marea Britanie au cerut răspunsuri la o nouă întrebare: de ce finanțăm extravagențele unor teoreticieni? Guvernul Statelor Unite, inclusiv DARPA, a retras finanțarea pentru proiectele de automatizare a traducerii. Companiile și-au schimbat prioritățile de la programe cronofage de cercetare generală a inteligenței artificiale la programe de amplitudine mai mică menite să rezolve probleme concrete. Dacă în primii ani de după atelierul de la Dartmouth a existat o perioadă de mari așteptări și optimism, de-

ceniile care au urmat acelor rapoarte au devenit cunoscute drept iarna inteligenței artificiale. Finanțarea a fost tăiată, studenții au preferat alte domenii de cercetare și avansurile au fost oprite.

Chiar și McCarthy și-a revizuit negativ predicțiile. „Oamenii pot face aceste lucruri foarte ușor pentru că așa sunt construiți", spunea McCarthy.[38] E mult mai greu să înțelegem care sunt modalitățile prin care recunoaștem vorbire – procesele fizice și cognitive care stau la baza recunoașterii limbajului. McCarthy dădea exemplul unei colivii pentru a ilustra provocările din domeniul inteligenței artificiale. Să spunem că v-aș cere să-mi construiți o colivie, și nu v-aș oferi niciun fel de alte date. Probabil că ați construi o cutie dotată cu un vârf, un fund, și niște laturi. Dacă v-aș mai da o informație – că pasărea e un pinguin – atunci poate nu i-ați mai pune un vârf. Prin urmare, prezența vârfului depinde de câțiva parametri: de informațiile pe care vi le dau, și de toate asocierile pe care vi le provoacă cuvântul „pasăre", cum ar fi că majoritatea păsările sunt zburătoare. Avem cu toții un anumit context și niște presupuneri implicite. Ca inteligența artificială să poată reacționa la fel ar presupune ca ea să dispună de mult mai multe informații și instrucțiuni explicite.[39] Iarna inteligenței artificiale urma să dureze trei decenii.[40]

Ce a urmat: învățarea jocurilor

Deși finanțarea s-a redus, mulți dintre cercetătorii de la Dartmouth au continuat să lucreze cu inteligența artificială – și au continuat să predea materia unor noi studenți. Între timp, legea lui Moore a rămas valabilă, iar calculatoarele au devenit din ce în ce mai puternice.

În anii 80, unii cercetători și-au dat seama cum ar putea vinde anumite trăsături ale inteligenței artificiale – iar acum exista suficientă putere de calcul și o rețea de cercetători în continuă creștere care constatau că din munca lor se pot face bani. Astfel, interesul publicului a fost repornit, și odată cu el, și investițiile în inteligența artificială au fost reluate. Japonia a anunțat în 1981 un plan pe 10 ani de dezvoltare a inteligenței artificiale intitulat „Generația a cincea". Guvernul Statelor Unite a reacționat înființând Microelectronics and Computer Technology Corporation, un consorțiu de cercetare menit să asigure competitivitatea economiei naționale. În Marea Britanie, fondurile care fuseseră tăiate în urma raportului lui James Lighthill au fost realocate. Între 1980 și 1988, industria inteligenței artificiale a crescut de la câteva milioane de dolari la câteva miliarde.

Calculatoarele mai rapide, încărcate cu memorie, puteau acum procesa mai eficient datele, iar accentul era pus pe reproducerea procesului decizional al experților, în loc de a construi mașini universale, ca fictivul computer HAL 9000. Aceste sisteme se axau în principal pe folosirea rețelelor neurale pentru sarcini mărunte, cum ar fi jocurile. În anii 90 și 2000 au avut loc niște succese rezonante. În 1994, un program denumit CHINOOK a jucat șase jocuri de dame împotriva campionului mondial Marlon Tinsley (toate partidele încheindu-se cu remize). CHINOOK a câștigat prin abandonul lui Tinsley.[41] În 1997, supercalculatorul IBM Deep Blue a câștigat în fața campionului mondial Garry Kasparov, care a cedat stresului unei finale de șase meciuri împotriva unui adversar aparent invincibil. În 2004, Ken Jennings a câștigat în ciuda probabilităților 74 de partide consecutive la *Riști și câștigi*,

stabilind un record mondial la acea vreme pentru sumele încasate dintr-un concurs televizat. Așa că atunci când a acceptat provocarea celor de la IBM de a juca împotriva lui Watson, în 2011, se simțea încrezător că va câștiga. Făcuse câteva cursuri de inteligență artificială și presupunea că tehnologia nu era suficient de avansată pentru a-și da seama de context, subtilități semantice și jocuri de cuvinte. Watson l-a umilit pe Jennings, care a început să-și piardă încrederea încă de la începutul jocului.

Ceea ce știam în 2011 e că inteligența artificială e mai bună decât oamenii la anumite sarcini de gândire pentru că poate accesa și procesa cantități uriașe de informații fără să cadă pradă stresului. Inteligența artificială poate defini stresul, dar nu are un sistem endocrin pe cap.

Cu toate acestea, jocul străvechi Go reprezenta apogeul cercetătorilor în inteligența artificială pentru că putea fi jucat doar pe baza unei strategii convenționale. Go este un joc străvechi apărut acum peste 3000 de ani, jucat cu reguli suficient de simple: doi jucători plasează pe rând pietre albe sau negre pe o grilă goală. Pietrele pot fi capturate atunci când sunt înconjurate de culoarea opusă, sau când nu mai există alte spații libere sau „libertăți." Obiectivul este să acoperiți teritoriu pe grila de joc, dar asta presupune aptitudini psihologice și o înțelegere precisă a stării de spirit a adversarului.

Dimensiunea tradițională a grilei de Go este de 19 cu 19 pătrățele. Spre deosebire de alte jocuri, precum șahul, pietrele de Go au valori egale. Există 181 de piese negre și 180 de piese albe (negrul mută întotdeauna primul, de unde diferența). În șah- unde piesele au diferite valori – jucătorul alb are la dispoziție 20 de mu-

tări, și cel negru 20 de mutări. După prima mutare, sunt 400 de poziții diferite ale tablei. În Go există însă 361 de prime mutări posibile, câte una la intersecția unei grile complet goale. După primele mutări ale fiecărui jucător, sunt 128.960 de mutări posibile. În total, există 10^{170} configurații posibile – ca să vă dați seama, e un număr mai mare decât numărul atomilor din universul cunoscut. Date fiind atâtea poziții și mișcări potențiale, nu există un manual, la fel ca în cazul șahului sau al damelor. Marii maeștri ai jocului de Go se bazează pe scenarii: dacă adversarul mută într-un anume punct, atunci care sunt rezultatele posibile, plauzibile și probabile, date fiind personalitatea, răbdarea și starea generală a adversarului?

La fel ca șahul, Go este un joc determinist cu informații perfecte, în care nu există niciun element ascuns sau de hazard. Pentru a câștiga, jucătorii trebuie să-și echilibreze emoțiile și trebuie să stăpânească la perfecție arta subtilității. În șah se pot anticipa eventualele mișcări probabile ale unui jucător; o tură se poate mișca doar vertical sau orizontal pe tablă. Numărul mișcărilor potențiale al acelei piese este deci limitat. Prin urmare, e mai ușor să înțelegeți cine câștigă un joc de șah cu mult timp înainte ca vreo piesă să fi fost capturată sau regele să fie pus în șah. În Go lucrurile nu stau așa. Uneori e nevoie de un mare maestru în Go care să-și dea seama ce se petrece și cine pare la un moment dat să câștige. Complexitatea jocului de Go i-a asigurat un public printre împărați, matematicieni și fizicieni. Acesta este motivul pentru care cercetătorii inteligenței artificiale au fost dintotdeauna fascinați de jocul de Go și au încercat să dezvolte programe care să joace Go.

Jocul de Go s-a dovedit dintotdeauna o nucă tare pentru cercetătorii inteligenței artificiale. Deși un calculator poate fi programat să învețe regulile jocului, cum rămâne cu regulile care ar permite înțelegerea caracteristicilor omenești ale adversarului? Nimeni nu mai construie un algoritm suficient de puternic care să poată face față complexităților acestui joc. În 1971, un program scris de informaticianul Jon Ryder a încercat o abordare tehnică, dar a pierdut în fața unui începător. În 1987, un program mai puternic denumit Nemesis a concurat pentru prima dată într-un turneu împotriva unui om. În 1994, programul Go Intellect se dovedea un jucător competent. Chiar și în ciuda unui handicap semnificativ, acesta a pierdut toate cele trei partide – împotriva unor copii. În toate aceste cazuri, calculatoarele făceau fie mișcări de neînțeles, fie jucau prea agresiv, sau calculau greșit strategia adversarului.

În cadrul acestor demersuri erau și câțiva cercetători care reluau ideea rețelelor neurale, o idee lansată de Marvin Minsky și Frank Rosenblatt la prima întâlnire de la Dartmouth. Cercetătorul cognitivist Geoff Hinton și informaticienii Yann Lecun și Yoshua Bengio credeau că rețelele neurale vor avea aplicații practice concrete – cum ar fi detectarea automată a fraudelor cu cărți de credit sau recunoașterea automată a caracterelor optice, pentru a citi documente sau cecuri – dar și că ele vor pune bazele viitorului inteligenței artificiale.

Hinton – profesor la universitatea din Toronto-a fost cel care și-a imaginat un nou tip de rețea neurală, alcătuită din diverse straturi care extrag diferite informații până când rețeaua recunoștea ce anume căuta. Singura modalitate de a oferi unui sis-

tem de inteligență artificială acest gen de cunoaștere, credea el, era de a dezvolta algoritmi de învățare care să le permită calculatoarelor să învețe pe cont propriu. În loc să le învățăm să facă o singură treabă simplă deosebit de bine, rețelele urmau să fie construite pentru a se învăța singure.

Aceste rețele neurale noi, profunde urmau să necesite un nou tip de învățare pentru mașini – învățarea profundă – pentru a antrena calculatoarele să efectueze sarcini omenești însă cu mai puțină (sau chiar nicio) supraveghere din partea oamenilor. Unul dintre avantajele imediate era faptul că proiectul putea fi extins la o scară mare. Într-o rețea neurală, câțiva neuroni fac câteva alegeri-numărul alegerilor posibile crește exponențial cu fiecare strat. Reformulând: oamenii învață individual, dar omenirea învață colectiv. Imaginați-vă o mare rețea neurală profundă, care învață ca un întreg – având posibilitatea de a-și spori cu timpul viteza, eficacitatea și randamentul.

Un alt avantaj îl reprezintă faptul că aceste sisteme ar fi libere să învețe pe cont propriu, fără să fie limitate de capacitățile cognitive și imaginația oamenilor. Creierul omenesc are niște praguri chimice și metabolice care limitează puterea de procesare a calculatoarelor umede din capul nostru. Nu putem evolua pe cont propriu în mod semnificativ, iar ritmul evoluției nu se potrivește cu aspirațiile noastre tehnologice actuale. Ni se oferea promisiunea că învățarea profundă ar accelera și însăși evoluția inteligenței, care ar presupune doar temporar un efort din partea oamenilor.

O rețea neurală profundă urma să primească un set elementar de parametri despre date din partea unei persoane-iar sistemul ar urma apoi să învețe pe cont propriu recunoscând tipare și folosind

mai multe straturi de procesare. Pentru cercetători, atracția învățării profunde este că mașinile sunt programate să ia decizii în mod imprevizibil. Gândirea într-o manieră pe care noi oamenii nu am putut-o concepe – sau pune în practică – e esențială atunci când încercăm să rezolvăm marile probleme, care nu au avut niciodată soluții clare.

Comunitatea inteligenței artificiale a considerat că rețelele neurale profunde sunt niște elucubrații ale unui cercetător excentric. Îndoiala lor s-a acutizat atunci când și-au dat seama că din cauză că procesele de învățare profundă se petrec în paralel, cercetătorii inteligenței artificiale nu le-ar putea observa în timp real. Cineva ar trebui să construiască sistemul, apoi să aibă încredere că lua deciziile corecte.

A câștiga și a pierde

Hinton a continuat să lucreze, organizând ateliere pe tema acestui concept cu studenții săi, dar și cu Lecun și Bengio, și a publicat articole, începând cu 2006. În anul 2009, laboratorul lui Hinton aplica rețelele neurale profunde pentru recunoașterea vocală, și dintr-o întâlnire întâmplătoare cu Li Deng, un cercetător de la Microsoft, a rezultat o etapă pilot a acestei tehnologii. Deng, un specialist chinez în învățare profundă, era unul dintre pionierii recunoașteriii vocale folosind învățarea profundă la scară largă. Tehnica intrase în teste la Google în 2010. Doar doi ani mai târziu, rețelele neurale profunde erau folosite în produsele comerciale. Dacă folosiți Google Voice și serviciile de transcriere vocală, ele se bazează pe învățare profundă, iar tehnica a devenit baza tuturor

asistenților digitali pe care-i folosim astăzi. Siri, Google și Alexa de la Amazon sunt toate exemple bazate pe învățarea profundă. Comunitatea interdisciplinară de cercetători în inteligență artificială a crescut semnificativ de la vara din Dartmouth. Aceste trei practici esențiale au rămas însă valabile-colaborarea dintre companiile mari din domeniul tehnologiei și cercetătorii din mediul academic; succesul comercial ca promotor al progreselor în domeniul inteligenței artificiale și omogenitatea rețelei de cercetători.

Toate aceste progrese americane nu au trecut neobservate la Beijing. China are acum un propriu ecosistem incipient, dar în creștere, iar guvernul central sprijină cercetătorii care vor să-și publice lucrările. Numărul de articole dedicate inteligenței artificiale publicate de cercetători chinezi a crescut cu peste 100% între 2010 și 2017.[42] La drept vorbind, articolele și brevetele nu înseamnă automat că aceste studii vor conduce și la o întrebuințare comună a acestei tehnologii, însă statisticile ne arată cât de tulburați au fost liderii chinezi de avansul occidentalilor- mai ales când vine vorba de jocul de Go.

În ianuarie 2014, Google a început să investească sume considerabile în inteligența artificială, printre care se numără o sumă de peste 500 de milioane de dolari pentru achiziția unei firme noi înființate în domeniul învățării profunde denumită DeepMind, fondată de neurobiologul Demis Hassabis, un fost mare jucător de șah în copilărie, cercetătorul în domeniul învățării mașinilor Shane Legg și antreprenorul Mustafa Suleyman. Faptul că dezvoltaseră un program denumit AlphaGo a constituit o parte din atracția acestei echipe.

Într-o perioadă de câteva luni au fost gata să testeze programul AlphaGo împotriva unui jucător uman. S-a programat un meci între

DeepMind și Fan Hui, un jucător chinez profesionist de Go, și unul dintre cei mai puternici maeștri din Europa. Deoarece a juca Go pe calculator nu e tocmai la fel cu a juca pe o tablă fizică, s-a decis ca unul dintre inginerii DeepMind să mute pe tablă în locul calculatorului și să-i transmită calculatorului mișcările acestuia.

Înaintea jocului, Toby Manning, unul dintre capii British Go Association, a jucat o rundă de test împotriva lui AlphaGo – și a pierdut la o diferență de 17 puncte. Manning a făcut niște greșeli, dar și programul a greșit. Acestuia i-a trecut un gând bizar prin cap: dacă AlphaGo jucase conservator? Era cu putință ca programul să fi jucat doar suficient de agresiv încât să-l înfrângă pe Manning, în loc să-l strivească ?

Jucătorii s-au așezat la o masă, Fan Hui purtând o cămașă în dungi și o jachetă de piele maro, Manning în centru și inginerul de cealaltă parte. Jocul a început. Hui a deschis o sticlă de apă și s-a uitat la tablă. Jucând cu piesele de culoare neagră, trebuia să înceapă. În primlee 50 de mișcări jocul a fost liniștit – Hui încerca în mod evident să sondeze punctele forte și cele slabe ale lui AlphaGo. Unul dintre primele semne: programul nu juca agresiv decât dacă rămânea în urmă. A fost un meci strâns. AlphaGo a câștigat o victorie foarte strânsă, de doar 1,5 puncte.

Hui s-a folosit de această informație când a început a doua partidă. Dacă AlphaGo nu avea de gând să joace agresiv, atunci Hui s-a decis că va lupta de la început. AlphaGo a început să mute mai repede. Hui a amintit că poate avea nevoie de mai mult timp ca să se gândească la mutări. La mutarea 147, Hui a încercat să împiedice ca AlphaGo să obțină un teritoriu mare în mijlocul tablei dar mutarea a dat greș, și a fost forțat să renunțe.

În jocul al treilea, Hui a mutat mai agresiv, și AlphaGo a procedat la fel. Undeva pe la jumătatea meciului, Hui a făcut o greșeală catastrofală pe care AlphaGo a speculat-o, apoi o altă greșeală mare care a condus la finalul jocului. Plin de frustrare, Hui a trebuit să se scuze și a plecat să se plimbe pentru a-și putea redobândi calmul și a termina partida. Din nou, stresul a doborât o minte strălucită-în vreme ce inteligența artificială și-a văzut fără scrupule de obiectiv.

AlphaGo – un program bazat pe inteligență artificială- a bătut cu 5-0 un jucător profesionist de Go. A reușit să câștige analizând mult mai puține mutări decât computerul DeepBlue al IBM. Când AlphaGo a bătut un jucător uman, acesta nu știa că joacă un joc, sau ce înseamnă un joc, sau de ce oamenilor le face plăcere să joace jocuri.

Hanjin Lee, un jucător bine cotat de Go din Korea, a analizat jocul dintre cei doi. „Impresia mea generală e că AlphaGo părea un jucător mai bun decât Fan, dar nu mi-am putut da seama cât de mare e acea diferență...poate că devine mai puternic atunci când joacă cu un adversar mai puternic."[43]

Accentul pus pe jocuri-a bate un om într-o confruntare directă – definește succesul folosind un set relativ redus de parametri. Ajungem astfel la întrebare filosofică pentru epoca modernă a inteligenței artificiale. Pentru ca sistemele inteligenței artificiale să câștige – pentru a îndeplini obiectivele pe care le-am definit- e nevoie ca oamenii să piardă, atât dintr-o perspectivă trivială, cât și una mai profundă?

Alpha Go a continuat să joace, depășind orice adversar ca un maer maestru și demoralizând comunitatea jucătorilor profesioniști de Go. După ce a câștigat împotriva campionului mondial cu 3-0, DeepMind a anunțat retragerea din competiție a sistemului, spunând că echipa se va concentra asupra unor noi provocări.[44] Echipa a încercat apoi să caute o variantă prin care AlphaGo să treacă de la un sistem puternic care putea fi antrenat să-i bată pe cei mai buni jucători la un sistem care ar putea învăța singur să ajungă la același nivel de expertiză fără să recurgă la ajutorul oamenilor.

Prima variantă a AlphaGo se baza pe intervenția oamenilor și un set inițial de date format din 100.000 de jocuri de Go pentru ca sistemul să poată învăța să joace. Următoarea generație a sistemului a fost construită să învețe de la zero. La fel ca un jucător începător, această variantă – denumită AlphaGo Zero- va trebui să învețe pe cont propriu de la zero, fără să dispună de variante din care să poată alege sau să aibă măcar o definiție a acțiunilor pe care le presupun mutările. Sistemul nu ar lua doar decizii – rezultatul unor calcule care pot fi programate în mod explicit – ci ar face alegeri, ceea ce necesită o judecată.[45] Aceasta presupunea că designerii DeepMind aveau o putere extraordinară, chiar dacă nu își dau seama de asta. Zero avea să învețe condițiile, valorile și motivările pentru deciziile și alegerile pe care le face în timpul jocului de la programatori.

Zero a concurat împotriva sa, ajustând și perfecționând de unul singur procesele decizionale. Fiecare joc începea cu câteva mișcări aleatorii, și cu fiecare câștig, Zero își actualiza sistemul și juca apoi luând în calcul ceea ce învățase. Zero a avut nevoie de

doar 70 de ore de joc pentru a ajunge la aceeași putere pe care o avea AlphaGo când câștigase împotriva celor mai buni jucători din lume.[46]

După care s-a produs ceva interesant. Echipa DeepMind a aplicat tehnica unei instalări separate a lui AlphaGo Zero folosind o rețea mai mare și i-a permis să se antreneze și să joace singură time de 40 de zile. Programul nu doar că a reușit să capete toate cunoștințele despre Go adunate de oameni de-a lungul istoriei, ci a reușit să învingă varianta avansată a AlphaGo în 90% din jocuri-asta înseamnă că Zero devenise un student mai bun decât cei mai mari maeștri de go și un profesor mai bun decât cei care l-au antrenat, și nu înțelegem în întregime ce a făcut pentru a deveni atât de inteligent.[47] Cam cât de inteligent credeți că e? Ei bine, puterea unui jucător de Go se măsoară pe baza unui index denumit Elo, care stabilește probabilitatea de a pierde sau a câștiga pe baza performanțelor anterioare. Marii maeștri și campionii mondiali au tendința să aibă un scor în jur de 3500. Zero are un scor de peste 5000. Prin comparație, aceste genii, campionii mondiali joacă ca niște amatori, și este statistic improbabil ca orice jucător uman să poată vreodată învinge acest sistem bazat pe inteligența artificială.

Știm care este condiția care a permis acest tip de învățare. Nefolosind vreo expertiză umană sau date, creatorii Zero au îndepărtat constrângerile omenești puse inteligenței artificiale. Se pare că influența oamenilor ar fi făcut sistemul mai puțin performant. Reușita a constat în crearea unui sistem care avea capacitatea de a gândi într-o manieră nouă și de a face propriile alegeri.[48] Sistemul a reprezentat un salt brusc, neașteptat, care prefigurează un viitor în care sistemele de inteligență artificială ar putea analiza radio-

grafiile pacienților de cancer, evalua datele climatologice și analiza sărăcia într-o metodă neomenească – putând eventual să conducă la progrese pe care cercetătorii umani să nu le fi putut niciodată obține pe cont propriu.

În timp ce Zero juca împotriva sa, a descoperit strategii de Go pe care oamenii le-au dezvoltat de-a lungul a mii de ani – ceea ce înseamnă că a învățat să gândească la fel ca oamenii care l-au conceput. La început făcea aceleași greșeli, detecta aceleași tipare și variații și tipare și întâmpina aceleași obstacole ca noi. Odată ce Zero a devenit suficient de puternic, a abandonat mișcările pe care le-ar fi făcut oamenii și și-a dezvoltat propriile preferințe.[49] Odată ce Zero a început să devină autonom, el și-a dezvoltat strategii creative pe care nimeni nu le mai văzuse, sugerând că poate mașinile *gândesc deja* în moduri care sunt atât recognoscibile, cât și străine de noi.

Zero a mai dovedit și că algoritmii sunt capabili să învețe ajutorul oamenilor, și că intervenția oamenilor e cea care înfrâna dezvoltarea inteligenței artificiale. Aceasta înseamnă că am putea încredința mașinilor problemele pe care noi înșine nu le putem prevedea sau rezolva.

În decembrie 2017, echipa DeepMind a publicat un articol în care arăta că Zero a devenit capabil să învețe – nu doar Go, dar și alte informații. Zero a început să joace alte jocuri, cum ar fi șah și shogi (un joc japonez asemănător cu șahul), care sunt mai puțin complexe dar necesită totuși creativitate și strategie. Zero învăța însă mult mai repede ca înainte. El a reușit să dezvolte o putere supra-omenească, de neînțeles, în mai puțin de 24 de ore de joc. Hassabis a început apoi să aplice tehnicile pe care le-a dezvoltat în

echipa sa pentru proiectul Zero pentru a construi o „mașină de învățat cu scop general", un set de algoritmi care se pot adapta și care imită sistemele noastre biologice și pot fi antrenate. În loc să umple sistemele inteligenței artificiale cu cantități uriașe de informații și instrucțiuni referitoare la felul în care pot fi accesate, echipa vă învăța mașinile cum să învețe. Spre deosebire de oameni care pot să obosească, să se plictisească sau să-și piardă atenția atunci când studiază, mașinile vor urmări cu orice preț un obiectiv.

Acesta a reprezentat un moment de cotitură în lunga istorie a inteligenței artificiale din mai multe motive. Mai întâi, sistemul se comporta de o manieră imprevizibilă, luând decizii care nu păreau raționale pentru creatorii lui. Sistemul a reușit să înfrângă un adversar uman fără ca să se poată înțelege sau reproduce complet cum a putut face asta. Apariția lui Zero prefigurează un viitor în care inteligența artificială și-ar putea construi propriile rețele neurale și să obțină o cunoaștere pe care nu am putea s-o înțelegem. În al doilea rând, apariția lui a cimentat cele două parcursuri paralele pe care le urmează evoluția inteligenței artificiale în prezent: alarmată, China alocă bani și oameni pentru a-și face propriile produse mai competitive, în timp ce în Statele Unite ne așteptăm să vedem produse fantastice bazate pe inteligența artificială și care să fie disponibile în curând. Viabilitatea rețelelor neurale profunde și a învățării profunde explică spatele frenezia din zilele noastre legate de inteligența artificială-ca să nu mai amintim de creșterea spectaculoasă a finanțării în Statele Unite și a programelor naționale ale Chinei pentru viitor.

Ca o companie din cadrul Alphabet (compania mamă a Google), DeepMind are 700 de angajați, iar o parte dintre ei au misiu-

nea de a dezvolta cât mai repede produse care să poată fi comercializate. În martie, divizia de servicii cloud a Google a anunțat că vinde un serviciu de citire cu voce tare a unui text bazat pe tehnologia DeepMind cu 16 dolari per 1 milion de caractere de text procesate.[50] Unul dintre anunțurile de senzație de la conferința Google I/O din 2018 a fost Duplex, un asistent vocal care va efectua automat apeluri în numele clienților și va discuta cu personalul pentru a face rezervări la restaurant sau programări la coafor, cu tot cu „mm"-uri și „ah"-uri. La baza acelui produs stă WaveNet, un program de generare software prin inteligență artificială, parte a portofoliului DeepMind.[51]

Între timp, cercetătorii din domeniul inteligenței artificiale din altă divizia a Alphabet denumită Google Brain au dezvăluit că dețineau o inteligență artificială capabilă să genereze propriile inteligențe artificiale. (Ați înțeles?) Sistemul, denumit AutoML, automatizează crearea modelelor de învățare automată printr-o tehnică denumită consolidarea învățării. AutoML funcționează ca un soi de „părinte" – o rețea neurală profundă care decide să creeze rețele de inteligență artificială „copii" pentru anumite sarcini concrete, țintite. Fără să i se spună să facă asta, AutoML a generat un copil denumit NASNet și l-a învățat să recunoască obiecte cum ar fi mașinile, semafoarele, oamenii sau gențile în videoclipuri. Neîmpovărat de ego, îndoieli sau lipsa încrederii- trăsături pe care chiar și cei mai sclipitori informaticieni le au – NASNet a reușit o acuratețe de 82,7% în recunoaștere imaginilor. Cu alte cuvinte, sistemul a reușit o performanță mai bună decât cea a programatorilor umani-inclusiv a celor care au creat sistemul părinte.[52]

Echipele de programatori care creează aceste sisteme menite să ia decizii și să facă alegeri sunt conduse de bărbați. Grupul e ușor mai diversificat decât cercetătorii de la Dartmouth din cauza unei evoluții semnificative ulterioare-intrarea Chinei în joc. În ultimii ani, China a devenit un centru important de cercetare în inteligența artificială, datorită unor demersuri colosale de finanțare ale guvernului la universități dar și companii ca Baidu, Alibaba și Tencent.

Baidu a reuși o performanță chiar mai mare decât a lui Zero: să transfere aptitudini dintr-un domeniu în altul. Pentru oameni e o misiune simplă, însă nu și pentru inteligența artificială. Baidu a urmărit să depășească acest obstacol învățând o rețea neurală profundă să navigheze într-o lume virtuală 2D folosind doar limbajul natural, așa cum vorbesc părinții cu copiii. Agentul inteligenței artificiale Baidu primea comenzi cum ar fi „te rog du-te către măr" sau „poți să te deplasezi către grila dintre măr și banană?" apoi primea recompense pentru acțiunile corecte. Pare ușor, însă gândiți-vă care sunt implicațiile: la finalul experimentului nu doar că inteligența artificială a Baidu devenise capabilă să înțeleagă un limbaj despre care nu știuse nimic, dar sistemul a și învățat ce este o grilă bi-dimensională, că se poate deplasa de-a lungul ei, cum să se deplaseze, că există mere și banane, și cum să le poată diferenția.

La începutul acestui capitol, am lansat patru întrebări: *Pot mașinile să gândească? Ce ar însemna ca o mașină să „gândească"? Ce înseamnă pentru voi, dragi cititori, să gândiți? Cum ați putea să*

vă dați seama că gândiți gânduri originale? Acum după ce ați aflat istoria îndelungată a acestor întrebări, micul grup de oameni care au pus bazele inteligenței artificiale și principalele practici care mai sunt folosite, aș dori să vă ofer câteva răspunsuri.

Da, mașinile pot gândi. Capacitatea de a trece un test conversațional, cum ar fi testul Turing sau recent apăruta Schemă Winograd, propusă de Hector Levesque în 2011 și care se concentrează asupra raționamentelor elementare, provocând inteligența artificială să răspundă la o întrebare care conține pronume ambigue, nu măsoară neapărat capacitatea unui sistem de inteligență artificială în alte domenii, ci dovedește doar că acest sistem poate gândi într-un cadru lingvistic, asemenea oamenilor.[53] Suntem cu toții de acord că Einstein era un geniu, chiar dacă metodele acceptate prin care se măsura inteligența în vremea respectivă – cum ar fi examenele – indicau un alt rezultat. Einstein gândea într-o manieră pe care profesorii lui nu o puteau înțelege- ei au presupus desigur că nu era inteligent. De fapt însă, la vremea respectivă nu exista o modalitate relevantă de a măsura inteligența lui Einstein. Același lucru este valabil și pentru inteligența artificială.

Mașinile care gândesc pot lua decizii și alege într-un mod care să aibă impact în lumea reală, iar pentru a face asta au nevoie de un scop și un obiectiv. În cele din urmă, și-ar putea dezvolta o judecată. Acestea sunt atribute ale sufletului, conform filosofilor și teologilor. Fiecare suflet este o manifestare a viziunii și voinței divine, fiind creat și înzestrat de un creator. Și mașinile care gândesc au creatori, ei sunt noii zei ai inteligenței artificiale, în majoritate bărbați care locuiesc cu preponderență în Statele Unite, Europa occidentală și China și care au legături cu cei nouă titani. Sufletul

inteligenței artificiale reprezintă o manifestare a viziunii și voinței lor pentru viitor.

Și da, pentru a concluziona, mașinile care gândesc sunt capabile de gândire originală. După ce învață prin experiență, ar putea constata că există o soluție diferită, sau că se impune o nouă clasificare. Inteligența artificială nu trebuie să inventeze o nouă formă de artă pentru a ne demonstra creativitatea.

Aceasta înseamnă că există în aceste mașini o minte. Această minte e încă tânără și în curs de formare, și e probabil ca ea să evolueze pe un parcurs pe care noi să nu-l înțelegem. În capitolul următor vom vorbi despre elementele acestei minți, valorile celor Nouă Titani și consecințele sociale, politice și economice neprevăzute ale acestei revoluții a inteligenței artificiale.

CAPITOLUL DOI

LUMEA INSULARĂ A TRIBURILOR INTELIGENȚEI ARTIFICIALE

Demersurile seculare de a construi o mașină capabilă să gândească nu au dat roade decât recent. Deși aceste mașini ar putea lăsa impresia că „gândesc", trebuie să înțelegem de la capul locului că ele nu gândesc cu siguranță la fel ca *noi*.

Viitorul inteligenței artificiale este construit de o minoritate relativ redusă de oameni din cadrul unor grupuri mici, atomizate. Repet, eu cred că acești oameni au intenții bune. Date fiind însă aceste grupuri insulare care colaborează strâns unele cu celelalte, prejudecățile lor inconștiente și o oarecare lipsă de perspectivă vor avea tendința să devină cu timpul noi sisteme de credință și comportamente acceptate. Ceea ce în trecut ar fi putut părea neobișnuit, chiar greșit, devine un nou normal cotidian, iar *această* gândire este programată în aceste mașini.

Cei care lucrează în domeniul inteligenței artificiale fac parte dintr-un soi de trib. Ei sunt niște oameni care trăiesc și lucrează în America de nord și China. Sunt oameni care au absolvit aceleași universități, și respectă aceleași reguli sociale. Triburile sunt foarte omogene. Vorbim de oameni bine educați și înstăriți. În mare parte triburile se compun din bărbați, conduși aproape în unani-

mitate de către bărbați-președinții, membrii consiliilor de administrație și directorii sunt aproape toți bărbați. Omogenitatea reprezintă o problemă și în China, unde membrii tribului sunt cu preponderență chinezi.

Problema triburilor reprezintă și unul dintre motivele care le conferă putere. În grupurile insulare, prejudecățile cognitive se amplifică și consolidează, instalându-se în subconștient. Prejudecățile cognitive sunt un substitut al gândirii raționale, care este mai înceată și presupune un consum mai mare de energie. Cu cât un trib devine mai interconectat și mai respectat, cu atât gândirea de grup și comportamentul grupului par mai normale. Veți constata mai târziu că merită să rețineți această concluzie.

Care este activitatea triburilor din domeniul inteligenței artificiale? Aceste triburi construiesc sisteme de inteligența artificială îngustă, capabilă să efectueze o singură sarcină la un nivel comparabil sau superior cu al oamenilor. Implementările comerciale ale aplicațiilor de inteligență artificială îngustă -prin extensie, triburile însele- iau deja decizii în numele nostru în căsuțele noastre de e-mail, când căutăm pe internet, când facem poze cu telefonul, când conducem mașina sau când solicităm un card de credit sau un împrumut. Triburile lucrează deja la următoarea etapă: sisteme de inteligență artificială generală care vor îndeplini sarcini cognitive mai vaste pentru că sunt mașini create să gândească, la fel ca noi. Cine sunt însă acești „noi" care servesc drept model acestor sisteme? Ce fel de valori, idealuri și perspective asupra lumii sunt transmise sistemelor?

Pe scurt, nu ale voastre, și nici ale mele. Inteligența artificială e concepută după chipul și asemănarea tribului, acordând priori-

tate valorilor, idealurilor și perspectivelor asupra lumii ale celor care au creat-o. Ea începe însă să gândească și pe cont propriu.

Liderii tribali

Tribul inteligenței artificiale are un strigăt de luptă cunoscut, antrenant: *dă greș repede, și dă greș des.* O altă variantă – „mișcă-te repede și distruge lucruri" era motto-ul oficial al Facebook până nu demult. Ideea de a comite greșeli și a accepta eșecul se află într-un contrast radical cu filosofia corporațiilor uriașe ale Americii, care evită riscurile și se mișcă cu viteza melcului, constituind un obiectiv lăudabil. Tehnologiile complicate precum inteligența artificială presupun experimente și ocazia de a da greș în mod repetat, cu scopul de a nimeri varianta corectă. Există, însă, și un revers al medaliei. Această mantră face parte dintr-o ideologie larg răspândită între cei Nouă Titani: *mai întâi construiește, cere-ți iertare mai târziu.*

În ultima vreme am tot auzit scuze. Facebook și-a cerut scuze pentru rezultatul relației cu Cambridge Analytica. În timpul desfășurării scandalului, Facebook a anunțat în luna septembrie 2018 ca un atac a vulnerabilizat datele personale a peste 50 de milioane de utilizatori, una dintre cele mai însemnate breșe de securitate din istoria digitală. Se pare însă că directorii au decis să nu anunțe imediat utilizatorii. [54] La doar o lună după aceea, Facebook a anunțat lansarea serviciului Portal, un ecran de video-conferințe care să rivalizeze cu Echo Show al Amazon, renunțând la promisiunile legate de securitatea datelor personale pe care le făcuse anterior. La început, Facebook spusese că nu va folosi serviciul Portal

pentru a colecta date personale în vederea țintirii anunțurilor publicitare către utilizatori. După ce jurnaliștii au cerut informații suplimentare, compania s-a văzut nevoită să ofere o lămurire bizară: deși Portal nu va folosi datele voastre pentru a afișa reclame, datele strânse în timp ce folosiți serviciul-persoanele apelate, melodiile ascultate pe Spotify- ar putea fi folosite pentru a vă oferi reclame Facebook prin alte servicii și rețele.[55]

În luna aprilie 2016, Jeff Dean, șeful de proiect al Google Brain scria că în timpul unei sesiuni de întrebări pe Reddit, compania exclusese femeile și persoanele de culoare. Nu era un gest premeditat ci o omisiune involuntară, și cred cu tărie că nu a fost voit, însă organizatorii pur și simplu nu s-au gând cum ar putea diversifica publicul sesiunii.

Dean spunea că pune preț pe diversitate și că Google va trebui să-și îmbunătățească practicile.[56]

> "Unul dintre lucrurile care chiar mi-a plăcut la programul nostru Brain Residency era că rezidenții provin dintr-o gamă largă de profiluri, domenii de expertiză (avem de exemplu fizicieni, matematicieni, biologi, neurobiologi, ingineri electricieni dar și informaticieni) și diferă între ei și în alte moduri, contribuind la demersurile noastre de cercetare", scria Dean. „Din experiența mea, ori de câte ori aduci laolaltă oameni cu expertize, perspective diferite etc vei ajunge să reușești mult mai mult decât ai fi putut pe cont propriu pentru că nu există un singur om înzestrat complet cu aptitudinile și perspectiva necesară pentru un asemenea exercițiu."[57]

Google a publicat în luna iunie 2018 un raport al diversității care cuprindea pentru prima oară date despre numărul de angajați în funcție de categorie. În cadrul raportului, Google spunea că la nivel global forța lor de muncă se compune în proporție de 69,1% din bărbați. În Statele Unite, doar 2,5% dintre angajați sunt negri, în vreme ce 3,6% sunt hispanici sau latino-americani. Cu toate afirmațiile îndrăznețe ale Google privind nevoia de a avea diversitate, aceste procentaje -deja scăzute- nu demonstrau vreo schimbare față de anii precedenți, în 2014 proporția negrilor fiind de 2% în vreme ce hispanicii și latino-americanii reprezentau 3% din total.[58]

Spre meritul lor, Google au lansat în ultimii ani o inițiativă împotriva prejudecăților inconștiente cu ateliere și cursuri menite să ajute angajații companiei să învețe mai multe despre stereotipurile sociale și atitudinile generale care s-ar fi putut forma inconștient privind genul, rasa, aspectul, vârsta, educația, politica sau bunăstarea. Unii angajați ai Google sunt de părere că aceste cursuri au mai degrabă un caracter de umplutură, iar o angajată de culoare a explicat că în cadrul cursului s-a vorbit cu precădere despre „relații interpersonale și sentimente rănite în loc să se discute despre discriminare și inegalitate, ceea ce transmite mesajul către angajați că diversitatea e doar ,o rubrică de bifat."[59]

Cu toate acestea, în timp ce avea loc acest curs, Google recompensa comportamentele nesănătoase ale conducerii. Andy Rubin, creatorul emblematicului sistem mobil de operare Android, fusese invitat să demisioneze după ce o angajată susținuse că acesta ar fi obligat-o să întrețină relații sexuale orale. Google i-a plătit lui Rubin suma de 90 de milioane de dolari pentru a de-

misiona-în tranșe lunare de 2,5 milioane de dolari în primii 2 ani și 1,5 milioane lunar în următorii doi ani. Directorul diviziei de cercetare și dezvoltare X, Richard DeVaul, a hărțuit sexual o femeie în cursul unui interviu de angajare, spunându-i că el și soția lui au o căsătorie fără angajamente, insistând apoi să mângâie acea candidată pe spate la un festival tehno. În mod evident, ea nu a fost angajată. I s-a cerut să prezinte scuze dar nu să demisioneze. Un vice președinte care se ocupa de divizia de căutare a Google a avut probleme atunci când o angajată l-a acuzat că ar fi pipăit-o-acuzația a fost considerată credibilă, așa că a fost concediat însă cu un pachet de compensare de mai multe milioane de dolari. Între 2016 și 2018, Google a renunțat la 13 directori din motive de hărțuire sexuală.[60]

Această reacție surprinde impactul șters pe care îl au multe programe de conștientizare în rândul firmelor din domeniul tehnologiei și al companiilor care le finanțează. Motivul este următorul: deși oamenii ar putea să fie mai conștienți de propriile prejudecăți după curs, ei nu devin neapărat motivați sau stimulați să-și modifice comportamentul.

Atunci când vorbim de lipsa diversității în rândul informaticienilor, conversația pendulează de regulă între gen și rasă. Există însă și alte dimensiuni cărora nu li se acordă la fel de multă atenție, ca de pildă ideologia politică și religia. Într-o analiză a Stanford Graduate School of Business efectuată în 2017, bazată pe răspunsurile a peste 600 de lideri și antreprenori, se arată că membrii tribului se consideră în proporție covârșitoare progresiști care susțin Partidul Democrat. În perioada alegerilor din 2016, membrii tribului au sprijinit-o în proporție covârșitoare pe Hillary Clinton.

Membrii tribului sunt de acord cu impozitarea mai mare a persoanelor înstărite, sprijină avorturile, se opun pedepsei cu moartea, vor un regim mai strict pentru permisele de port armă și susțin legalizarea căsătoriilor homosexuale.[61]

Afirmația că persoanele aflate la conducerea Google, Apple, Amazon, Facebook, Microsoft sau IBM nu reprezintă în mod fidel o imagine a *tuturor* americanilor s-ar putea aplica și altor companii din alte domenii. Diferența este însă că aceste companii sunt cele care dezvoltă sisteme autonome de luare a deciziilor care sunt menite să reprezinte toate interesele noastre. Criticile provin nu doar din partea femeilor sau a persoanelor de culoare, ci și din partea unor grupuri neașteptate: conservatorii și susținătorii Partidului Republican. În luna mai 2018, Comitetul Național Republican i-a trimis o scrisoare lui Mark Zuckerberg acuzând Facebook de discriminare împotriva americanilor de orientare conservatoare reclamând, printre altele, că: „În ultimii ani au apărut îngrijorări legate de suprimarea mesajelor conservatoare pe Facebook...inclusiv cenzurarea articolelor cu tentă conservatoare... Ne alarmează numeroasele aserțiuni conform cărora Facebook blochează conținutul publicat de grupuri și jurnaliști conservatori."[62] Scrisoarea, semnată de Ronna Mc Daniel, președintele Comitetului Național Republicat și Brad Parscale, șeful de campanie pentru realegerea președintelui Trump în 2020, solicită transparență în ce privește felul în care algoritmii Facebook stabilesc utilizatorii care sunt expuși la publicitate politică și analizarea prejudecăților împotriva conținutului și liderilor conservatori.

Problema e că McDaniel și Parscale nu greșesc. În timpul alegerilor febrile din 2016, angajații Facebook au manipulat cu bună

știință platforma pentru a exclude știrile cu conținut conservator – chiar și prin articole împotriva lui Clinton și care începuseră să capete propria notorietate. Câțiva dintre „curatorii de știri" de la Facebook, așa cum sunt aceștia cunoscuți, au afirmat că li s-a cerut să „introducă" anumite articole în flux, chiar dacă acestea nu începuseră să fie urmărite. Curatorii au împiedicat și afișarea articolelor favorabile unor candidați ai Partidului Republican, precum Rand Paul. Echipa de curatori a Facebook era alcătuită dintr-un număr mic de jurnaliști care absolviseră facultăți private de elită de pe coasta de est, și, ca să fim sinceri, acest aspect susține explicațiile oferite de conservatori vreme de decenii.

În luna august 2018, peste 100 de angajați Facebook au folosit un grup intern pentru a dezavua „cultura politică monocromă, intolerantă față de alte păreri." Unul dintre inginerii seniori ai Facebook, Brian Amerige, scria: „Pretindem că suntem deschiși tuturor perspectivelor, dar atacăm – adesea în haită – pe oricine prezintă o perspectivă care pare să se opună ideologiei de stânga."[63]

A vorbi despre diversitate – a cere iertare, a promite că vei face mai bine-nu e totuna cu a rezolva problema diversității în bazele de date, algoritmii și structurile care compun sistemul inteligenței artificiale. Când vorbele nu conduc la fapte, rezultă un ecosistem de sisteme și produse care reflectă o anumită prejudecată împotriva oamenilor. În 2016, un robot de securitate controlat de inteligența artificială a intrat intenționat într-un copil în vârstă de 16 luni într-un centru comercial din Silicon Valley. [64] Sistemele de inteligență artificială aflate la baza jocului video *Elite: Dangerous* au creat o serie de super-arme pe care creatorii nu le-au imaginat niciodată, făcând ravagii în joc și distrugând orice pro-

grese făcute de jucătorii în carne și oase.[65] Există nenumărate probleme legate de siguranța inteligenței artificiale, dintre care unele sunt mari și evidente: mașinile care conduc singure au trecut deja pe roșu și chiar au ucis pietoni în câteva rânduri. Aplicațiile folosite de poliție etichetează greșit chipurile suspecților în mod constant, băgând oameni nevinovați în arest. Există nenumărate probleme necunoscute pe care nu le băgăm în seamă pentru că încă nu ne-au afectat personal.

O echipă cu adevărat diversă ar avea o singură trăsătură comună: talentul. Nu ar exista o concentrare a unui anumit gen, sau a unei anumite rase sau etnicități. Ar exista reprezentanți ai mai multor perspective politice și religioase. Omogenitatea triburilor inteligenței artificiale constituie o problemă a celor Nouă Titani, însă ea nu începe aici. Problema începe în universități, acolo unde se formează triburile.

Triburile apar în cadrul unor medii sociale concentrate în care toată lumea are un scop sau un obiectiv comun, folosește același limbaj și muncește într-un ritm relativ similar. În cadrul tribului, grupurile de oameni își dezvoltă un simț al valorilor și scopurilor împărtășite. Triburile se formează în locuri precum unitățile militare, gărzile din facultatea de medicină, bucătăriile restaurantelor cu stele Michelin și organizațiile studențești. Membrii lor trec împreună prin bune și rele, reușite și eșecuri, fericire și tristețe.

Pentru a prelua un exemplu dintr-un domeniu îndepărtat de inteligența artificială, în anii '70 și '80, Sam Kinison, Andrew Dice Clay, Jim Carrey, Marc Maron, Robin Williams și Richard Pryor au locuit împreună într-o casă situată pe Cresthill Road la numărul 8420, la doi pași de ceea ce avea să devină legendarul club Comedy

Store din Los Angeles. Erau doar niște tineri care locuiau într-o casă, încercând să obțină reprezentații pe scenă într-o perioadă în care Bob Hope apărea la televizor cu replici ca „Nu mă gândesc de două ori la femei. În primul gând spun tot."[66] Acest trib respingea total acest gen de umor, pe care generația anterioară îl perfecționase cu meticulozitate. Valorile lor erau complet diferite: distrugerea tabuurilor, combaterea nedreptăților sociale și relatarea unor povești cu lux de amănunte care aveau tendința să prezinte o imagine destul de neplăcută a celor aflați în public. Ei repetau între ei. Își țineau de urât după un eșec pe scenă. Experimentau și învățau unii de la alții. Acest trib de comici revoluționari, inteligenți a pus bazele viitorului în divertismentul american.[67] Acest grup de bărbați exercită la nivel colectiv o influență chiar și în zilele noastre.

Inteligența artificială a trecut cumva printr-o perioadă asemănătoare de transformări radicale din cauza unui trib din zilele noastre care împărtășește aceleași valori, idei și obiective. Aceștia erau cei trei pionieri ai învățării profunde amintiți anterior – Geoff Hinton, Yann Lecun și Yoshua Bengio – un soi de Sam Kinison și Richard Pryors ai lumii inteligenței artificiale în epoca de pionierat a rețelelor neurale profunde. Lecun a studiat cu Hinto la Universitatea din Toronto unde s-au pus bazele unui mic grup de cercetători printre care se numără și Yoshua Bengio. Ei petreceau perioade lungi de timp împreună, testând idei, teorii și construind următoarea generație a inteligenței artificiale. „Era o comunitate foarte mică de oameni care anticipaseră că rețelele neurale vor deveni într-o bună zi importante," spunea Lecun. „Aveam nevoie de o zonă sigură în care să putem organiza ateliere și întâlniri pentru a ne dezvolta conceptele înainte de a le publica."[68]

Legăturile puternice ale tribului se formează atunci când oameni care lucrează îndeaproape sărbătoresc reușitele și suferă în momente grele împreună. Ei își dezvoltă un set de experiențe comune, prin care ajung la un limbaj comun, care are ca rezultat un set comun de idei, comportamente și obiective. Acesta este motivul pentru care atât de multe firme nou înființate, mișcări politice sau simboluri culturale au un început comun: câțiva prieteni care împart un dormitor, o casă sau un garaj și lucrează intens la proiecte asemănătoare.

Dacă epicentrele inteligenței artificiale moderne sunt Silicon Valley, Beijingul, Hangzhou și Shenzhen, facultățile sunt esența triburilor inteligenței artificiale. Există doar câteva centre importante. Printre ele, în Statele Unite se numără Carnegie Mellon, Georgia Institute of Technology, Stanford, UC Berkeley, University of Washington, Harvard, Cornell, Duke, MIT, Boston University, McGill University și Université de Montréal. Aceste universități găzduiesc grupuri de cercetare academică, cu puternice legături cu industria IT.

Triburile respectă în general regulile și ritualurile, așa că haideți să vedem în ce constă inițierea. Procesul începe cu o educație temeinică.

În universitățile din America de nord, accentul se pune pe competențele brute-cum ar fi stăpânirea limbajelor de programare R și Python, competențe în procesarea limbajului natural, statistică aplicată și cunoașterea teoriei jocurilor, bioinformatică și vedere digitală. Cursurile din afara tribului, cum ar fi filosofia minții, literatura feministă islamică sau istoria colonialismului nu sunt văzute cu ochi buni. Dacă încercăm să construim o mașină care să

gândească și să fie capabilă să gândească la fel ca oamenii, pare contraintuitiv să excludem anumite aspecte ale condiției umane. În prezent acest gen de cursuri este exclus intenționat din programa de bază și e dificil ca ele să fie acceptate chiar și ca cursuri opționale.

Tribul pretinde aptitudini, și în patru ani de facultate sunt multe de învățat. La Stanford, spre exemplu, studenții trebuie să facă 50 de ore de matematică, științe și inginerie, pe lângă 15 ore de informatică. Deși în cadrul diplomei se oferă și un curs de etică, acesta este unul dintre cele cinci cursuri opționale care pot fi alese pentru a obține creditele necesare.[69] Carnegie Mellon a lansat un program nou de diplomă în inteligență artificială în 2018, dând facultății un nou imbold și ocazia să creeze un program modern de la zero. Ritualurile și regulile tribului s-au impus însă, iar aptitudinile brute sunt cele care contează. Deși obținerea diplomei presupune un curs de etică și câteva cursuri socio-umane și de artă, acestea se concentrează în mare parte pe neuorobiologie (psihologie cognitivă, memorie și cogniție vizuală), ceea ce e de înțeles dată fiind legătura dintre inteligența artificială și mintea omenească. Nu există cursuri obligatorii care să învețe studenții să detecteze prejudecățile în seturile de date, cum să aplice filosofia în procesul de luare a deciziei sau să le predea etica egalității. Aceste cursuri nu recunosc în mod formal că diversitatea socială și socioeconomică sunt la fel de importante pentru o comunitate ca și biodiversitatea.

Aptitudinile sunt predate experimental- studenții din domeniul inteligenței artificiale nu-și petrec timpul citind din cărți. Pentru a putea învăța, ei au nevoie de baze de date lexicale, fototeci și

rețele neurale. La un moment dat, una dintre cele mai populare rețele neurale de la universități se numea Word2vec, fiind construită de echipa Google Brain. Această rețea era un sistem cu două straturi care procesa textul, transformând cuvintele în numere pe care inteligența artificială să le poată înțelege.[70] Spre exemplu, sistemul a învățat că „bărbatul e față de rege așa cum femeie e față de regină." Baza de date a hotărât însă și că „tată e față de doctor așa cum mamă e față de asistentă" și „bărbat e față de programator așa cum femeie e față de casnică."[71] Sistemul la care erau expuși studenți avea propriile prejudecăți. Dacă cineva ar fi vrut să analizeze implicațiile profunde ale codului cu prejudecăți, nu existau cursurile care să evidențieze aceste aspecte.

În 2017 și 2018 unele universități au oferit câteva cursuri noi de etică, ca reacție la provocările inteligenței artificiale. Departamentul Berkman Klein Center de la Harvard și MIT Media Lab oferă împreună un curs nou de etică și reglementare a inteligenței artificiale.[72] Cursul și sesiunile sunt excelente,[73] dar nu fac parte din programa de informatică de bază iar din această cauză materia nu are ocazia să se răspândească și în alte părți ale programei.

E drept că etica e o cerință a tuturor universităților care predau inteligența artificială-cerința se regăsește în standardele de acreditare. Pentru a fi acreditate de Accreditation Board for Engineering and Technology, programele de studii în informatică trebuie să arate că studenții „înțeleg problemele și răspunderile de ordin profesional, etic, juridic, social și de siguranță" și „au capacitatea de a analiza impactul local și global al calculatoarelor asupra oamenilor, organizațiilor și societății." Vă pot spune însă din experiență că evaluarea și măsurarea acestei cerințe este cel mult

subiectivă și deosebit de greu de pus în practică, mai ales în absența unor cursuri obligatorii pentru toți studenții. Eu sunt membră a Accrediting Council on Education in Journalism and Mass Communications. Programa de studii a diplomelor de jurnalism și comunicare pune accentul pe discipline socio-umane, cum ar fi scrisul, reportajele și producția media. Cu toate acestea, facultățile întâmpină în mod regulat probleme în ce privește îndeplinirea standardelor noastre legate de probleme și răspunderi sociale, inclusiv în ce privește diversitatea. Facultățile pot să fie acreditate fără să respecte standardele în ce privește diversitatea - și acest aspect nu e specific doar comisiei de acreditare din care fac parte. Fără o impunere mai strictă a standardelor și fără demersuri concrete în cadrul universităților, cum ar putea o programă încărcată de studiu ca cea a inteligenței artificiale să rezolve problema?

Facultatea e destul de solicitantă, iar stimulentele oferite noilor angajați de către cei Nouă Titani sunt substanțiale. Deși cursurile opționale privind literatura africană sau etica funcționarilor publici ar lărgi cu siguranță perspectivele asupra lumii ale lucrătorilor din inteligența artificială, există o presiune puternică de dezvoltare a ecosistemului. Tribul preferă aptitudinile legate de muncă astfel încât atunci când proaspeții absolvenți intră pe piața muncii să poată fi gata de lucru și să fie membri productivi ai echipelor din care fac parte. De fapt, chiar cursurile care i-ar putea ajuta pe cercetătorii din domeniul inteligenței artificiale să se gândească mai atent la întreaga omenire le-ar putea diminua șansele de angajare. Cei Nouă Titani folosesc programe bazate pe inteligența artificială pentru a filtra CV-urile, căutând anumite cu-

vinte cheie care descriu aptitudinile. Un portofoliu mai diversificat de cursuri din afara materiei standard ar reprezenta o anomalie sau ar face candidatul mai greu de recrutat.

Faptul că inteligența artificială evaluează CV-urile dovedește că prejudecățile nu sunt doar o chestiune de rasă sau gen. Există prejudecăți și împotriva filosofiei, literaturii, fizicii sau economiei comportamentale deoarece candidații cu multe cursuri opționale în afara domeniului de bază al inteligenței artificiale au tendința să fie cotați mai slab. Sistemul de angajare al tribului, creat să automatizeze munca dificilă de a evalua mii de CV-uri ar putea să excludă acești candidați care au un profil mai divers și mai de dorit de studii.

Liderii din domeniul academic vor spune că sunt deschiși la ideea unui curs obligatoriu de etică, chiar și dacă tribul nu ar dori o programă mai vastă de studiu. (Și tribul nu vrea acest lucru.) Adăugarea unor cursuri la fel de solicitante din domeniul socio-uman, cum ar fi literatura comparată sau religiile lumii, ar forța scoaterea din program a anumitor cursuri prin care studenții ar dobândi aptitudinile legate de inteligența artificială. Studenții s-ar revolta că sunt forțați să facă și cursuri aparent inutile, în vreme ce companiile pretind absolvenților aptitudini de vârf. Dată fiind concurența acerbă de a obține cei mai buni și mai inteligenți studenți, de ce și-ar dori vreuna dintre aceste universități prestigioase, cum ar fi Carnegie Mellon sau Stanford, să modifice o rețetă de succes?

Domeniul tehnologiei se schimbă mult mai rapid decât mediul universitar. Un singur curs obligatoriu de etică -gândit în mod special în funcție de nevoile studenților din domeniul inteligenței

artificiale- nu va fi suficient dacă materia de studiu nu e de actualitate dar mai ales dacă ea nu se regăsește și în alte părți ale programei. Dacă programa nu se poate schimba, ce se poate face în privința profesorilor? Ar putea ei să primească răspunderea de a rezolva problema? La scară mare, acest rezultat este improbabil. Profesorii nu sunt încurajați să-și modifice cursurile pentru a lega materia de întrebări legate de valorile tehnologice, economice și sociale. Dacă ar trebui să facă asta ar pierde un timp prețios. Cursurile lor ar putea deveni mai puțin atrăgătoare pentru studenți. Universitățile vor să demonstreze că studenții lor se pot angaja după absolvire, iar angajatorii vor absolvenți cu aptitudini puternice. Cei Nouă Titani au parteneriate cu aceste universități, care depind de finanțarea și resursele lor. Cu toate acestea, a venit momentul ca întrebările dificile *-cine e proprietarul tău?-* să fie puse și dezbătute în mediul universitar, înainte ca studenții să devină membri ai unor echipe aflate în mod constant sub presiunea termenelor limită și a țintelor de încasări.

Dacă universitățile sunt locul unde se nasc triburile inteligenței artificiale, e ușor de văzut de ce în acest domeniu diversitatea este mult mai scăzută față de alte profesii. Directorii companiilor de tehnologii se pripesc să arate cu degetul înspre universități, dând vina pe ele pentru diversitatea scăzută a forței de muncă. Nu e o acuzație complet falsă. Triburile inteligenței artificiale se formează în laboratoare și săli de clasă, în timp ce studenții colaborează la diferite teme și proiecte. Acești profesori, laboratoarele și conducerea departamentelor de inteligență artificială din facultăți sunt în proporție covârșitoare compuse din bărbați și cu o diversitate scăzută.

În cadrul universităților, doctoranzii îndeplinesc trei funcții: să colaboreze în activitatea de cercetare, să predea studenților și să ducă mai departe domeniul. Femeile beneficiază de doar 23% din doctoratele în informatică și doar 28% din doctoratele în matematică și statistică, conform datelor recente ale National Center for Education Statistics.[74] Femeile dispar pe parcursul academic: doctorandele nu sunt promovate în roluri de conducere sau ca profesori titulari în egală măsură cu bărbații. Nu ar trebui deci să surprindă pe nimeni că femeile au reprezentat doar 18% din absolvenții de informatică din ultimii ani – procentaj chiar mai scăzut decât cel din 1985, de 37%.[75] Candidații negri sau latino-americani sunt și ei foarte absenți – constituind doar 3%, respectiv 1% din totalul doctoranzilor.[76]

Pe măsură ce tribul crește, expansiunea se produce în interiorul unei bule, provocând unele comportamente inacceptabile. Cercetătoarele din domeniul inteligenței artificiale au trebuit să se confrunte cu hărțuire sexuală, glume nepotrivite și un comportament în general execrabil al colegilor lor. Pe măsură ce acest comportament devine considerat normal, el se transferă odată cu tribul dinspre universitate înspre forța de muncă. Prin urmare, nu e atât o problemă a parcursului de instruire, cât o problemă de *oameni*. Triburile inteligenței artificiale sprijină pur și simplu o cultură din care femeile și anumite minorități-negrii sau latino-americanii – sunt excluși.

Un inginer Google a trimis o notă de informare devenită celebră în 2017 prin care susținea că femeile sunt biologic inferioare când vine vorba de programare. Directorul general al Google, Sundar Pichai, a reacționat în cele din urmă concediindu-l pe angajat, însă a afirmat și că „putem dezbate o bună parte din acea notă de informare."[77] Culturile ostile față de membrii din afara tri-

bului amplifică acest efect, prin urmare forța de muncă devine și mai puțin diversă. Pe măsură ce inteligența artificială progresează, construind sisteme care gândesc pentru și alături de oameni, segmente întregi din populație sunt excluse.

Nu vreau să spun că în universități nu lucrează și femei sau oameni de culoare. Directorul faimosului Computer Science and Artificial Intelligence Lab de la MIT este Daniela Rus, o femeie care a fost bursieră MacArthur, printre multe alte realizări profesionale și academice. Kate Crawford este profesor emerit la New York University și conduce un institut nou înființat care studiază efectele inteligenței artificiale asupra societății. Există și femei și oameni de culoare care fac o treabă grozavă în domeniul inteligenței artificiale – numărul lor este însă foarte redus.

Dacă tribul urmărește să ofere inteligenței artificiale o gândire mai „umanistă", mulți oameni sunt însă excluși din acest proces. Fei-Fei Li, care conduce laboratorul de inteligență artificială de la Stanford și este cercetător principal în domeniul inteligenței artificiale și învățării mașinilor la Google Cloud spune că

> „ca femeie, ca educator, ca femeie de culoare, ca mamă, sunt din ce în ce mai îngrijorată. Inteligența artificială urmează să aducă cele mai mari schimbări din istoria omenirii, iar nouă ne lipsește o generație întreagă de informaticieni și lideri cu diversitate... Dacă nu reușim să integrăm femeile și persoanele de culoare – în rândul informaticienilor care lucrează pe bune- vom prejudicia sistemele. Dacă vom încerca peste zece sau douăzeci de ani să revenim asupra acestei situații va fi mult mai dificil, dacă nu chiar aproape imposibil."[78]

Triburile Chinei: BAT

Baidu, Alibaba și Tencent, cunoscute împreună sub numele de BAT reprezintă partea chineză a celor Nouă Titani. Tribul inteligenței artificiale din Republica Populară Chineză operează cu reguli și ritualuri diferite, pe baza unor finanțări, controale și politici industriale ale guvernului care sunt menite să-i împingă pe membrii BAT spre viitor. Această structură bine finanțată și organizată implementează planurile la nivel național pentru implementarea viitoare a inteligenței artificiale, iar guvernul deține o putere extraordinară în cadrul structurii. Această epocă este echivalentul cursei spațiale pentru China, iar noi suntem pentru ei ce reprezenta Sputnik pentru misiunea Apollo. Poate că am fost primii care am ajuns pe orbită, însă China a pus la bătaie fondul suveran de investiții, sistemul educațional, cetățenii și mândria națională în cursa pentru inteligența artificială.

Și triburile inteligenței artificiale din China își au rădăcinile în sistemul universitar, unde accentul pe aptitudini și aplicațiile comerciale este și mai pronunțat. Deoarece China este interesată să sporească cât mai repede aptitudinile forței de muncă, problemele de diversitate nu sunt tocmai comparabile cu cele ale Occidentului, deși ele există. Genul nu joacă un rol atât de mare, astfel încât femeile sunt mai bine reprezentate. Cu toate acestea, cursurile sunt predate în chineză, o limbă dificil de stăpânit pentru străini. Se exclud astfel persoanele nevorbitoare de chineză, creându-se însă și un avantaj concurențial special, deoarece studenții chinezi cunosc în general limba engleză și pot studia în mai multe universități.

În China, instruirea în inteligență artificială începe înainte ca studenții să intre la facultate. În 2017, consiliul de stat al Chinei a cerut includerea unor cunoștințe de bază în inteligența artificială în programa școlară, ceea ce înseamnă că elevii chinezi încep să se pregătească din timpul *școlii primarei*. Există în prezent un manual oficial editat sub egida guvernului care prezintă istoria și elementele de bază ale inteligenței artificiale. Până în 2018, 40% dintre licee au luat parte la un program pilot cu un curs obligatoriu de inteligență artificială,[79] și odată ce vor fi disponibili mai mulți profesori și alte școli vor fi incluse în program. Nu va dura prea mult pentru că ministerul educației din China a lansat un program de instruire în inteligență artificială în universități, cu peste 500 de profesori și 5000 de studenți de la universitățile de vârf din China.[80]

Membrii BAT contribuie la revoluția educației din China, oferind instrumentele folosite în școli și universități, vânzând pe piață produsele pe care consumatorii le folosesc în copilărie și la maturitate, angajând absolvenții și făcând schimb de experiență cu guvernul. Dacă nu ați călătorit sau trăit în China în ultimii 10 ani, poate că Baidu, Alibaba și Tencent nu sunt niște nume prea cunoscute. Toate cele trei companii au fost înființate în aceeași perioadă, folosind ca model companiile din domeniul tehnologiei care existau în acea perioadă.

Baidu a fost înființată în cadrul unui picnic de vară din Silicon Valley-una dintre acele ocazii exclusiviste în care membrii tribului inteligenței artificiale se adună și beau bere. Trei bărbați, toți pe la vreo 30 de ani, se plângeau de cât de puțin evoluează motoarele de căutare. John Wu, la vremea respectivă directorul echipei de motoare de căutare a Yahoo și Robin Li, inginer la Infoseek, cre-

deau că motoarele de căutare au înainte un viitor strălucit. Văzuseră deja ce rezultate avea o firmă nou înființată-Google – și credeau că pot face ceva asemănător în China. Împreună cu biochimistul Eric Xu, cei trei au înființat compania Baidu.[81]

Compania a recrutat angajați din centrele universitare de cercetare în inteligență artificială din America de nord și China. Compania se pricepea îndeosebi să atragă cercetători cu aptitudini în domeniul învățării profunde. În anul 2012, compania i-a făcut o ofertă lui Andrew Ng, un cercetător de vază de la divizia Brain a lui Google. el a crescut în Hong Kong și Singapore și a trecut prin centrele importante în inteligență artificială: diplomă de licență la Carnegie Mellon, diplomă de master la MIT, doctorat la UC Berkeley, apoi o carieră didactică la Stanford. Ng le-a atras atenția celor de la Baidu pentru că lucra la un proiect revoluționar al unei rețele neurale profunde la Google.

Echipa lui Ng construise o rețea de 1000 de calculatoare care s-a învățat singură să recunoască pisicile în videoclipurile de pe Youtube. Era un sistem remarcabil. Fără să i se fi spus niciodată în mod explicit ce este o pisică, inteligența artificială a vizionat milioane de ore de videoclipuri la întâmplare, a învățat să recunoască obiecte, și-a dat seama că anumite obiecte sunt pisici apoi a învățat ce este o pisică. Toate pe cont propriu, fără ca oamenii să aibă vreo influență. La puțin timp după aceea, Ng s-a angajat la Baidu ca cercetător principal. (Acest lucru înseamnă că în ADN-ul Baidu se regăsesc obligatoriu nucleotide din cursurile de inteligență artificială de la Carnegie Mellon, MIT și UC Berkeley.)

În ziua de azi, Baidu nu e doar un motor de căutare. Ng a contribuit la platforma de conversații bazată pe inteligență artificială

(DuerOS), un asistent digital și programe autonome de condus mașina, precum și alte aplicații ale inteligenței artificiale care au permis Baidu să înregistreze progrese în domeniu cu mult înaintea Google. Baidu are o capitalizare de piață de 88 de miliarde de dolari și e al doilea motor de căutare din lume după Google – o mare realizare având în vedere că nu este folosit în afara Chinei. La fel ca Google, Baidu dezvoltă o gamă de aparate casnice inteligente, cum ar fi un robot domestic care îmbină recunoașterea vocală cu cea facială. Compania a anunțat o platformă deschisă pentru conducerea autonomă a vehiculelor intitulată Apollo, având speranța ca statutul deschis al codului sursă să provoace o înflorire a ecosistemului limitrof. Baidu colaborează cu peste 100 de parteneri, printre care se numără constructorii de automobile Ford și Daimler, fabricanții de chipuri NVIDIA și Intel și firme de cartografie precum TomTom. Baidu a încheiat un parteneriat cu compania californiană Access Services pentru a pune la dispoziția persoanelor cu probleme de mobilitate mașini autonome. Baidu a încheiat un parteneriat și cu divizia Azure Cloud a Microsoft pentru a permite partenerilor Apollo din afara Chinei să proceseze cantități uriașe de date despre mașini.[82] Ar trebui să știți și că Baidu a inaugurat în ultimii ani un laborator nou de cercetare în inteligență artificială în parteneriat cu guvernul chinez – iar liderii laboratorului sunt nomenclaturiști cu experiență în programele armatei.[83]

A-ul din BAT corespunde Alibaba Group, o platformă uriașă care joacă rol de intermediar între cumpărători și vânzători printr-o rețea gigantică de siteuri, nu printr-o platformă unică. Alibaba a fost înființat în 1999 de către Jack Ma, un fost profesor care trăiește la aproximativ 150 de km la sud-vest de Shanghai,

care și-a dorit să creeze o variantă hibrid între Amazon și eBay pentru piața din China. Ma nu avea cunoștințe informatice, așa că a înființat compania împreună cu un coleg de facultate. După doar 20 de ani, Alibaba are o capitalizare de piață de peste 511 miliarde de dolari.

Printre siteurile deținute de Alibaba se numără Taobao, pe care atât cumpărătorii cât și vânzătorii nu plătesc nicio taxă pentru tranzacțiile efectuate. Taobao folosește în schimb un model, prin care taxează vânzătorii care doresc să obțină un loc mai bun pe motorul de căutare al siteului. (Această tactică imită o parte din modelul de afaceri al Google.) Alibaba a dezvoltat și sisteme de plată sigure, printre care se numără Alipay, cu o funcționalitate și caracteristici asemănătoare cu PayPal. Alibaba a lansat un sistem de plăți digitale „zâmbește ca să plătești" bazat pe inteligență artificială și recunoaștere facială, care permite consumatorului să plătească zâmbind în fața unei camere.

La fel ca Amazon, Alibaba are un difuzor inteligent – se numește Genie X1, e mai mic și îndesat de Alexa de la Amazon sau aparatele Home ale Google. Genie X1 folosește o tehnologie de recunoaștere a amprentei vocale bazată pe rețele neurale pentru a identifica utilizatorii, autentificându-i automat pentru a le permite să vizualizeze produse și să le cumpere. Peste 100.000 de astfel de aparate Alibaba sunt în curs de instalare în hotelurile Marriott din toată China.

Alibaba are o viziune de ansamblu asupra inteligenței artificiale, pe care o denumește „ET City Brain." Programul prelucrează cantități uriașe de date obținute local, de la camere și senzori inteligenți din oraș, la arhive guvernamentale și conturi de pe

rețele de socializare. Alibaba folosește această infrastructură de inteligență artificială pentru a genera modele predictive: pentru a îmbunătăți traficul, urbanismul, sănătatea publică și a stabili dacă urmează să se producă tulburări sociale. Sub conducerea lui Ma, Alibaba a intrat și în domeniul logisticii, găzduirii de videoclipuri online, al centrelor de stocare a datelor și tehnologiilor cloud, investind miliarde de dolari în diferite companii, urmărind să construiască un colos digital care să conecteze lumea de afaceri, casele, munca, orașele și guvernul. De fapt, înainte ca magazinul Amazon Go să fie inaugurat în Seattle, Alibaba lansase magazinul Hema, un concept automat, fără bani gheață, care combină un magazin alimentar, o piață și un serviciu de livrare la domiciliu.

Mai există o asemănare stranie care merită amintită aici. Spun „stranie" pentru că ea reprezintă și o contradicție. În anul 2016, Ma a cumpărat ziarul *South China Morning Post*, cel mai mare și mai influent cotidian din Hong Kong. Tranzacția a fost una importantă pentru că în China majoritatea publicațiilor sunt finanțate de stat iar acest ziar de limbă engleză a fost cunoscut pentru articolele critice la adresa guvernului chinez.[84] Când trăiam în Hong Kong, obișnuiam să beau cu mai mulți reporteri de la acest ziar, niște jurnaliști de investigații de primă mână. Prin achiziția acestui ziar, Ma își demonstra loialitatea față de partidul comunist. Cu trei ani în urmă, Jeff Bezos cumpărase *Washington Post*, o mutare care l-a transformat în cele din urmă într-un inamic al Casei Albe conduse de președintele Trump din cauza jurnalismului obiectiv al ziarului, analizei critice asupra politicilor acestei administrații și a demersurilor continue de combatere a propagandei.[85]

În ultimul rând, cel mai mare și în multe privințe cel mai influent membru al BAT este Tencent. Această companie a fost înființată în 1998 de doi bărbați, Ma Huateng și Zhang Zhidong. Aceștia au pornit de la un singur produs denumit OICQ. Dacă vi se pare că sună ușor cunoscut, asta e pentru că era o copie a serviciului de mesagerie ICQ. Cei doi au fost dați în judecată dar au continuat să lucreze la propria variantă a sistemului. În 2011, Tencent a lansat WeChat, care oferă nu doar un serviciu de mesagerie ci copiază și funcționalitățile și caracteristicile Facebook. Din momentul în care guvernul chinez a luat decizia de a bloca Facebook în China, WeChat a dobândit o poziție avantajoasă. Platforma a depășit granițele mediului academic, fiind folosită pentru recrutarea noilor angajați-și în multe alte scopuri.

WeChat are o cifră uluitoare de 1 miliard de utilizatori activi în fiecare lună și o poreclă „aplicația pentru orice." Asta pentru că dincolo de postările standardizate pe rețelele de socializare și mesagerie, WeChat este folosită pentru aproape orice în China, de la recrutarea noilor angajați în universități și mesagerie text la plăți virtuale și chiar măsuri de ordine publică. Peste 38.000 de spitale și clinici au conturi pe WeChat, și 60% dintre acestea le folosesc pentru gestionarea relației cu pacienții (de ex., programări și plăți).[86] Tencent este o companie care se sprijină – și se concentrează – pe inteligența artificială, considerând că ea „reprezintă o tehnologie fundamentală în toate produsele noastre diferite."[87] Nu e de mirare că sloganul oficial al companiei este „Inteligența artificială peste tot."

Facebook ar putea fi cea mai mare rețea de socializare din lume, însă tehnologia de care dispune Tencent este din multe

puncte de vedere superioară. Tencent a construit un asistent digital numit Xiaowei, un sistem mobil de plăți (TenPay), un serviciu cloud (Weiyun) și a lansat recent și un studio de film (Tencent Pictures). Laboratorul YouTu al Tencent obține rezultate de vârf în recunoașterea facială și a obiectelor, și transferând aceste tehnologii către peste 50 de alte departamente ale companiei. Tencent are și inițiative în domeniul sănătății, încheind parteneriate cu două companii de sănătate din Marea Britanie: Babylon Health, o firmă nou înființată din domeniul medicinei la distanță și Medopad, care folosește inteligența artificială pentru monitorizarea pacienților de la distanță. Tencent a investit masiv și în două companii promițătoare din Statele Unite, Atomwise și XtalPi care vizează aplicarea inteligenței artificiale în domeniul farmaceutic.

Tencent a devenit prima companie din Asia care a depășit în 2018 o valoare de piață de 550 de miliarde de dolari și a depășit Facebook, devenind cea mai valoroasă rețea de socializare din lume.[88] Ceea ce e și mai uimitor e că sub 20% din veniturile Tencent provin din publicitatea online, în vreme ce Facebook depinde în proporție de 98% de aceste venituri.[89]

Pepiniera BAT include centrele universitare din America de nord, iar compania se asigură că copiii primesc o educație în inteligență artificială începând cam de la vârsta la care învață să adune și să scadă.

Toate aceste aspecte nu ar fi atât de importante dacă grupul BAT nu s-ar bucura de un succes extraordinar- și dacă nu ar câștiga sume uriașe de bani. BAT produce atât de mult iar piața din China este atât de mare, încât triburile inteligenței artificiale din China au o putere uriașă- atât în China cât și în alte părți ale lumii. Co-

munitatea globală a inteligenței artificiale e atentă la China din cauza acestui capital și a cifrelor care sunt greu de ignorat.

Facebook ar putea avea 2 miliarde de utilizatori activi în fiecare lună, dar ei sunt împrăștiați în toată lumea. Miliardul de utilizatori WeChat se găsește în mare măsură într-o singură țară. Baidu are 665 de milioane de utilizatori care folosesc căutarea mobilă în 2017[90] – mai mult decât dublul numărului estimat de deținători de telefoane mobile în Statele Unite.[91] În același an, Amazon a înregistrat cele mai bune vânzări de sărbători din istorie. Pentru a putea înțelege, între ziua Recunoștinței și Cyber Monday,[b] clienții Amazon au comandat 140 de milioane de produse, cu o valoare totală de 6,59 miliarde de dolari. [92] Poate că pentru Amazon ar fi fost un record, însă nici nu se compară cu realizările Alibaba în China, unde în doar 25 de ore, Alibaba a vândut unui număr de 515 milioane de clienți în 2017, iar de ziua oamenilor singuri, un soi de Black Friday combinat cu premiile Oscar în China – a vândut de 25 de miliarde de dolari pe internet, procesând 812 milioane de comenzi *într-o singură zi*.[93] După orice standarde, China are cea mai mare piață digitală din lume: peste 1.000 de miliarde de dolari cheltuiți anual, peste un miliard de oameni pe internet, și 30 de miliarde de dolari investite speculativ în cele mai importante companii din domeniul tehnologic din lume.[94]

Investitorii chinezi au fost implicați în 7% până la 10% din toate finanțările companiilor nou înființate în domeniul tehnologic din Statele Unite în perioada 2012-2017- vorbim de sume sem-

b Ziua Recunoștinței se sărbătorește în Statele Unite în ultima zi de joi a lunii noiembrie, iar Cyber Monday este prima zi de luni de după Ziua Recunoștinței (n.t.)

nificative de bani care provin dintr-o singură regiune.[95] BAT au acum o prezență palpabilă în Seattle și Silicon Valley, cu propriile filiale, inclusiv birouri pe faimoasa stradă Sand Hill Road. În ultimii cinci ani, BAT a investit sume considerabile în companii ca Tesla, Uber, Lyft, Magic Leap (un fabricant de căști de realitate virtuală) și altele. Investițiile companiilor din grupul BAT sunt atrăgătoare nu doar pentru că se mișcă rapid și au mulți bani- ci și pentru că un acord cu grupul BAT presupune de regulă o portiță profitabilă către piața din China, unde altminteri ar fi imposibil de pătruns. Spre exemplu Zoloz, o companie mică din Kansas City care se ocupa cu recunoașterea facială, a fost cumpărată de Alibaba pentru suma de 100 de milioane de dolari în 2016, devenind apoi o componentă de bază a sistemului de plăți Alipay, obținând acces la sute de milioane de utilizatori fără a trebui să se confrunte cu legile draconice legate de intimitate din Europa, sau cu amenințarea unor potențiale procese în Statele Unite. Investiția nu a venit însă lipsită de compromisuri serioase. Investitorii chinezi nu așteaptă doar ca investiția să dea randament- ei cer și drepturile de proprietate intelectuală asupra tehnologiei.

În China, faptul de a pretinde să primești drepturile de proprietate intelectuală asupra tehnologiei în schimbul banilor nu reprezintă o particularitate culturală, sau o modalitate lacomă prin care anumiți investitori să câștige mai mult. Această abordare reprezintă un demers coordonat al guvernului. China are o viziune clară asupra felului în care vrea să domine global din punct de vedere economic, geopolitic și militar- și consideră că inteligența artificială este calea către acest obiectiv. În vederea acestui scop, menținerea unui control autoritar asupra informațiilor reprezintă

o problemă foarte importantă pentru conducătorii țării, așa că China a adoptat o politică industrială menită să transfere proprietatea intelectuală de la companiile din Statele Unite către cele chineze. Printre exemple se numără anumite seturi de date, algoritmi sau designul procesoarelor. Multe companii americane care speră să facă afaceri în China trebuie să promită mai întâi să ofere acces la propriile tehnologii. Au fost adoptate noi reglementări, prin care companiile străine trebuie să-și desfășoare activitatea de cercetare și dezvoltare în China și să stocheze datele folosite local. Stocarea locală a datelor este o cerință dură impusă companiilor din străinătate deoarece guvernul chinez ar putea să invoce autoritatea sa pentru a accesa oricând aceste date, ocolind metodele de criptare.

Beijingul ia în serios planificarea pe termen lung. Tradiția a început încă din vremea lui Mao Zedung, care a introdus primele cincinaluri începând cu anul 1953. (Președintele Xi a lansat cel de-al 13-lea cincinal în 2016).[96] Atât liderii politici cât și nomenclaturiștii sunt adepții prevederii strategice- făcând din China una dintre puținele țări de pe glob care are o strategie economică, politică, militară și socială amănunțită pentru următoarele decenii. Guvernul chinez are capacitatea unică de a implementa orice politică își dorește și de a face orice este necesar pentru a-și pune în aplicare strategia națională, inclusiv planul de a transforma până în 2030 China în „principalul centru de inovare în inteligența artificială din lume" și să creeze o industrie în valoare de 150 de miliarde de dolari. E puțin probabil ca acest plan să fie anulat de către un nou guvern, deoarece China a abrogat în luna martie 2018 limitele privind mandatele, permițând președintelui Xi Jinping să rămână la putere pe viață.

Sub Xi, China a înregistrat o consolidare impresionantă a puterii. El a întărit Partidul Comunist, a suprimat fluxul informațiilor și a instituit noi politici pentru accelerarea unei multitudini de planuri pe termen lung, de pe urma cărora se așteaptă să beneficieze în deceniul următor. La nivelul cel mai de sus al guvernului chinez, inteligența artificială ocupă o poziție centrală. Spre deosebire de fostul lider al Partidului Comunist Chinez, Deng Xiaoping, care guverna în baza principiului „să ne ascundem capacitățile și să tragem de timp", Xi e gata să arate lumii ce poate China să facă – și intenționează să preia inițiativa la nivel global.[97] Conducerea Chinei e orientată spre viitor și are de gând să-și pună în aplicare chiar acum planurile îndrăznețe, unite. Acest aspect oferă Chinei un avantaj incredibil față de Occident, iar membrilor grupului BAT le conferă o putere extraordinară.

Toate acestea se petrec într-o perioadă de creștere economică puternică a Chinei, care înregistrează o creștere extraordinară a clasei de mijloc. Până în 2022, mai bine de trei sferturi din populația urbană a Chinei va dispune de suficiente resurse pentru a face parte din clasa de mijloc. În 2000, doar 4% erau considerați a face parte din clasa de mijloc- vorbim de o creștere uriașă într-un interval scurt de timp. Slujbele mai bine plătite din domeniul informaticii, biotehnologiilor și serviciilor vor ridica o mare parte a acestui grup din pătura de mijloc în partea superioară a clasei de mijloc. Gospodăriile din China au un grad de îndatorare foarte redus. Deși e adevărat că peste tot în China există și sărăcie, generația actuală a copiilor din China are oportunități bune de a câștiga mai mult, a economisi mai mult și a cheltui mai mult decât părinții lor.[98] (În mod frapant, 70% dintre americani se consideră

parte a clasei de mijloc, dar datele Pew Research Center[99] arată că clasa de mijloc americană s-a micşorat în ultimele patru decenii – mai puţin de jumătate dintre americani câştigă suficient pentru a se putea încadra în această categorie.[100])

China a devenit o putere economică greu de ignorat. Compania Marriott a semnat un contract pentru instalarea a 100.000 de aparate inteligente Alibaba în hotelurile din China, însă când Beijingul a aflat că pe siteul companiei Hong Kong, Taiwan, Tibet şi Macau sunt trecute ca ţări diferite într-un chestionar trimis pe e-mail membrilor clubului de clienţi fideli, directorii Marriott au primit o somaţie imediată. Guvernul i-a comunicat companiei Marriott să-şi închidă toate siteurile şi aplicaţiile din China, iar compania s-a supus. Compania are un portofoliu de peste 240 de hoteluri şi staţiuni de lux deschise în toată China pentru a profita de expansiunea clasei de mijloc. Directorul general Arne Sorenson a fost nevoit să publice un comunicat uluitor pe siteul companiei: „Marriott International respectă şi sprijină suveranitatea şi integritatea teritorială a Chinei. Din nefericire, în ultima săptămână au avut loc două incidente care nu confirmă această abordare: mai întâi, printr-o etichetare incorectă a unor regiuni din China, inclusiv a Tibetului, ca fiind ţări într-un meniu dintr-un sondaj pe care-l trimitem clienţilor fideli, iar apoi prin faptul că unul dintre angajaţii noştri a apreciat o postare care sugera în mod incorect că am sprijini acest punct de vedere. Nimic nu e mai departe de adevăr: nu sprijinim pe nimeni care subminează suveranitatea şi integritatea teritorială a Chinei şi nu dorim în vreun fel să încurajăm sau să incităm aceşti oameni sau aceste grupuri. Recunoaştem gravitatea situaţiei şi ne cerem sincer scuze."[101]

China a devenit și o forță geopolitică prea puternică pentru a fi subminată. Ea pune presiune asupra guvernelor străine printr-un plan de politică externă menit să revitalizeze pentru secolul 21 fostul drum multimilenar al mătăsii. China cheltuie 150 de miliarde de dolari pe an în 68 de țări pentru repararea infrastructurii: drumuri, căi ferate, poduri, porturi. Prin aceste programe țările participante se vor înscrie din ce în ce mai mult pe orbita Beijingului, într-o perioadă în care America bate în retragere. În timp America e ocupată cu incertitudinile și haosul administrației Trump, președintele Xi a făcut din China un pol al stabilității. În absența Americii, Xi a început să umple golul de putere la nivel global.

Donald Trump, de pildă, a repetat pe Twitter în timpul campaniei afirmații care neagă încălzirea globală, inclusiv referitoare la o teorie bizară a conspirației conform căreia chinezii ar dori să ne saboteze economia.[102] Evident că nu există niciun sâmbure de adevăr în toată povestea asta. China a construit alianțe în ultimul deceniu, urmărind să reducă deșeurile de plastic, să obțină mai multă energie regenerabilă și să elimine poluarea industrială. China nu prea avea de ales: după decenii în care a fost inima industrială și groapa de gunoi a lumii, poluarea a provocat probleme de sănătate generalizate și reducerea speranței de viață în China. În 2017, guvernul a anunțat că China nu va mai importa deșeuri din alte țări, după ce , începând cu 1992, China a importat și procesat 106 milioane tone de deșeuri.[103] Fiindcă Statele Unite nu își fac planuri pe termen lung, nu avem la dispoziție o alternativă. În prezent nu avem la dispoziție alternative pentru depozitarea deșeurilor, ceea ce înseamnă că China forțează celelalte țări să nu

mai folosească bunuri care nu pot fi reciclate. China devine rapid un lider global în problema dezvoltării durabile, fiind suficient de puternică pentru a-și impune punctul de vedere.

În China, oamenii apreciază *chengyu*, niște aforisme din patru ideograme prin care transmit înțelepciunea. Unul care-mi vine acum în minte este 脱颖而出, care în traducere înseamnă „bobul se leapădă de pleavă și dă rod."[104] China își arată acum lumii foarte fățiș puterea de o manieră foarte evidentă.

Consolidarea puterii lui Xi, combinată cu puterea și ascensiunea economică a Chinei pun bazele înfloririi triburilor inteligenței artificiale, în special dat fiind demersul consolidat vertical în domeniul inteligenței artificiale. La marginea Beijingului se construiește un parc de cercetare în valoare de 2 miliarde de dolari care se va concentra pe învățare profundă, tehnologii cloud și biometrice și va include un laborator național de cercetare și dezvoltare. Nu doar că guvernul investește în grupul BAT ci îl și protejează de concurența globală. Guvernul chinez interzice accesul Google și Facebook, și a zădărnicit încercarea Amazon de a intra pe piață. Companiile din grupul BAT se află în centrul planului guvernamental de dezvoltare până în 2030, care se bazează puternic pe tehnologiile acestor companii: sistemele autonome de condus ale Baidu, tehnologia Internet of Things a Alibaba și sistemele comerciale interconectate și aplicațiile Tencent în domeniul interfețelor de comunicare și asistenței sanitare.

Iată de ce ar trebui să vă preocupe triburile inteligenței artificiale din China, indiferent în ce țară locuiți.

Economia Chinei crește într-un ritm alert, iar evoluția rapidă a inteligenței artificiale nu face decât să accelereze ascensiunea Chinei. La finalul anului 2017, analiza și modelele pe care le-am efectuat împreună cu echipa de la Future Today Institute au arătat că inteligența artificială are potențialul de a spori economia Chinei cu 28% până în 2035. Inteligența artificială – alimentată de numărul uriaș de cetățeni chinezi și datele lor, automatizarea la scară largă, mașinile care învață și eficientizarea cheltuielilor de capital – va stimula creșterea industriei, agriculturii, comerțului, serviciilor financiare, transporturilor, utilităților, sănătății și divertismentului (inclusiv a platformelor) din China. În momentul de față nu există o altă țară cu atât de multe date, atât de mulți oameni și atât de multe aparate electronice pe cap de locuitor. Nicio altă țară nu are șansa de a depăși Statele Unite în decurs de o generație. Nicio altă țară nu are un potențial mai mare de a influența ecosistemul și clima planetei – conducând fie la supraviețuire, fie la catastrofă- decât China. Nicio altă țară nu constituie o punte între lumea dezvoltată și cea în curs de dezvoltare în aceeași măsură ca China. Cu un regim comunist dar o economie puternică, China e atât un partener prea mare pentru a putea fi ignorat, un adversar politic cu puncte de vedere complet diferite în ce privește drepturile omului, dar și un coagulant al alianțelor globale. Prosperitatea economică aduce puterea. China se poziționează pentru a putea influența fluxurile financiare și comerciale mondiale. Prin această reașezare celelalte țări își pierd puterea și influența, iar democrația are de suferit în lume.

În al doilea rând, China se va folosi de progresele din domeniul inteligenței artificiale și de stimulente economice pentru a-și

moderniza armata, oferindu-i un avantaj asupra țărilor din Occident. Această tranziție a început deja ca parte a unui program național de supraveghere aeriană cu numele de cod „Porumbelul." Peste 30 de agenții militare și guvernamentale au lansat „păsări spion" care arată ca niște păsări albe care bat imitând aripile păsărilor. Dronele fac parte dintr-un program inspirat din natură menit să evite radarul și detecția.[105] Dronele filmează, iar un sistem de inteligență artificială caută tipare, recunoaște fețe și identifică anomalii. Cu toate acestea, deși păsările spion sună amenințător, ele sunt cel mai mic motiv de îngrijorare pe care l-ați putea avea.

La finalul anului 2017, un raport nepublicat al Pentagonului obținut de către reporterii Reuters avertiza că firmele chineze se sustrag supravegherii guvernamentale a Statelor Unite și accesează tehnologii secrete din domeniul inteligenței artificiale din Statele Unite care ar putea avea și aplicații militare cumpărând companii din Statele Unite. Armata Populară de Eliberare a Chinei investește puternic într-o gamă de proiecte și tehnologii din domeniul militar, în timp ce institutele de cercetare militare colaborează cu industria de apărare din China.[106]

China nu a mai luptat împotriva unei alte țări din 1979, cu ocazia războiului chino-vietnamez. S-ar părea că China nu are adversari militari de calibru – nu a fost ținta unor atacuri teroriste, nu are relații antagonice cu suspecții de serviciu (de ex., Rusia, Coreea de Nord) și nu are alți dușmani. De unde atunci această militarizare intensă?

Motivul este că, în viitor, războaiele vor fi purtate prin intermediul programelor, nu prin lupte la baionetă. Folosind tehnologiile inteligenței artificiale, o armată poate „câștiga" destabilizând

o economie, în loc să pustiască satele și orașele. Dată fiind această perspectivă și demersurile puternice ale Chinei în domeniul inteligenței artificiale, China are un avans periculos față de Occident.

Din punctul meu de vedere, ne-am dat seama prea târziu de acest lucru. Când m-am întâlnit cu oficialitățile de la Pentagon a fost dificil ca perspectiva nouă asupra războiului (programe versus luptă) să fie acceptată. De exemplu, în 2017, departamentul Apărării a înființat o echipă interdisciplinară de război algoritmic pentru a contribui la proiectul Maven – un sistem de vizualizare digitală și învățare profundă care recunoaște automat obiecte din imagini și videoclipuri. Echipa nu dispunea de resursele necesare în inteligență artificială așa că departamentul a încheiat un contract cu Google pentru a pune la punct sistemele care să analizeze imaginile filmate din drone. Angajaților Google nu li s-a spus însă că vor lua parte la un proiect militar, iar reacțiile au fost vehemente. Patru mii de angajați Google au semnat o petiție prin care își exprimau dezacordul față de proiectul Maven. Ei au cumpărat o pagină de publicitate în *New York Times*, și în cele din urmă zeci de angajați au demisionat.[107] Google a comunicat că nu va reînnoi contractul cu departamentul Apărării.

Și Amazon a fost luată în vizor din cauza unui contract cu Pentagonul în valoare de 10 miliarde de dolari. În luna octombrie 2018, Tom Cole, republican, membru al comisiei pentru buget din Oklahoma, și Steve Womack, republican din Arkansas, au acuzat Departamentul Apărării că a redactat caietul de sarcini în colaborare cu Amazon pentru a exclude alte companii. Aceasta nu a fost însă singura acuzație. Unii dintre angajații Amazon s-au scanda-

lizat la aflarea veştii că angajatorul lor lucrează pentru armata Statelor Unite, în vreme ce alţii nu au apreciat faptul că tehnologiile de recunoaştere facială sunt puse la dispoziţia forţelor de ordine. Ca răspuns, Jeff Bezos a declarat în faţa unei conferinţe de presă: „Dacă companiile mari din domeniul tehnologiei vor întoarce spatele Departamentului Apărării, ţara noastră va avea mari probleme."[108]

Dacă în Statele Unite giganţii tehnologici încearcă să jongleze între transparenţă şi securitatea naţională, relaţia dintre grupul BAT şi guvernul chinez se află la polul opus. Iată însă un exemplu îngrijorător: poziţia actuală a armatei americane este că oamenii trebuie să deţină controlul, indiferent cât de avansate devin inteligenţa artificială, sistemele autonome şi roboţii. Prin acest control ne asigurăm că nu lăsăm decizia de a folosi forţa letală în mâinile unui program. China are o abordare diferită.[109] General-locotenentul Liu Guozhi, şeful comisiei de tehnologie şi ştiinţă a armatei chineze ar fi avertizat „(noi) trebuie...să preluăm iniţiativa pentru a schimba paradigma."[110] Afirmaţia reprezintă o modalitate indirectă de a anunţa intenţia Chinei de a-şi reconstrui puterea militară.

În al treilea rând, dacă nu vă îngrijorează superioritatea economică şi militară a Chinei, perspectiva lor asupra intimităţii ar trebui să vă îngrijoreze. Din nou, de ce v-ar interesa atât timp cât nu locuiţi acolo? Pentru că guvernele autoritare se formează tot timpul, iar ele au tendinţa de a copia regimurile deja existente. Dată fiind creşterea la nivel global a naţionalismului, felul în care China se foloseşte de inteligenţa artificială ar putea deveni un model şi pentru alte state în anii care vor urma. Acest fapt ar putea dezechilibra pieţele, comerţul şi balanţa geopolitică de putere.

China se folosește de inteligența artificială în încercarea de a crea o populație mai obedientă, unul dintre cele mai vaste și mai insidioase experimente din istoria omenirii. Planul consiliului de stat pentru inteligență artificială până în 2030 explică faptul că inteligența artificială va „îmbunătăți semnificativ eficiența și nivelul guvernării sociale" și va fi folosită pentru a juca „un rol de înlocuit în menținerea eficientă a stabilității sociale."[111] Strategia se implementează prin intermediul sistemului chinezesc național de credite sociale, care, conform cartei de înființare a consiliului de stat, va „permite celor de încredere să se plimbe peste tot sub cer, iar cei discreditați abia vor face un singur pas."[112] Această idee a apărut încă din 1949, când Partidul Comunist a preluat puterea și a început să experimenteze diferite modalități de control social. În vremea regimului lui Mao Zedung din anii 50, controlul strict al societății a devenit norma: muncitorii au fost colectivizați și primeau calificative în funcție de productivitate. Oamenii se urmăreau unii pe alții iar ierarhia calificativelor stabilea cât de mult puteau accesa bunurile publice. Sistemul s-a descompus în vremea lui Mao, apoi din nou în anii 80 pentru că, după cum s-a dovedit, oamenii nu se pricep de minune să se evalueze reciproc, fiind motivați de propriile nevoi, nesiguranțe și prejudecăți.

În 1995, președintele Jiang Zemin și-a imaginat un sistem de coerciție socială bazat pe tehnologie -iar după 2005, guvernul chinez a lucrat pentru a construi și implementa un sistem de notare care funcționa automat.[113] În parteneriat cu Universitatea din Beijing, guvernul a înființat centrul de cercetare pentru credite din China, pentru a cerceta modalitățile de construire și implementare a unui sistem național de credite sociale bazat pe inteligența arti-

ficială. Iată deci o explicație parțială a insistențelor președintelui actual Xi pentru dezvoltarea inteligenței artificiale. Prin dezvoltarea acestui sistem nu s-ar împlini doar un ideal de la înființarea Partidului Comunist, ci implementarea lui ar reprezenta chiar un element care ar sprijini menținerea partidului la putere.

Un sistem algoritmic de credite sociale a dovedit deja că inteligența artificială funcționează în orașul Rongcheng. Cei 740.000 de locuitori adulți ai orașului primesc la înscrierea în sistem 1000 de puncte, iar funcție de comportamentul demonstrat punctajul poate fi scăzut sau crescut. O faptă „eroică" poate aduce unui locuitor 30 de puncte, în timp ce a trece pe culoarea roșie a semaforului atrage o penalizare de 5 puncte. Cetățenii sunt etichetați și împărțiți în diferite categorii, de la A+++ la D, iar alegerile pe care le pot face și capacitatea de a se deplasa sunt dictate de categoria din care fac parte. Cei care fac parte din categoria C ar putea constata că li se solicită o garanție dacă vor să închirieze o bicicletă, în timp ce cei din categoria A le pot închiria pe gratis timp de 90 de minute. Sistemul nu se aplică doar persoanelor fizice. În Rongcheng, și *companiile* primesc note în funcție de comportament- iar capacitatea lor de a opera depinde foarte mult de categoria din care fac parte.[114]

Străzile din Shanghai sunt împânzite de microfoane și camere inteligente conectate la un sistem bazat pe inteligență artificială. Șoferii care claxonează excesiv sunt amendați automat și primesc amenda prin serviciul WeChat al TenCent, în timp ce numele, fotografia și cartea de identitate sunt afișate pe panouri luminoase aflate în apropiere. Dacă un șofer își lasă mașina parcată mai mult de șapte minute, el primește automat amendă.[115] Nu e vorba doar

de amendă-șoferul este penalizat și în sistemul de credite sociale. După ce se înregistrează abateri repetate, persoana ar putea întâmpina probleme în a-și cumpăra bilete de avion sau a se angaja. Un episod popular al serialului *Black Mirror* ilustrează un viitor distopic similar. În Shanghai acest viitor a devenit deja realitate.

Supravegherea la nivel național se efectuează pe baza tehnologiilor puse la dispoziție de grupul BAT, care e la rândul său stimulat prin intermediul diferitelor politici administrative și industriale ale guvernului chinez. Serviciul de credit Zhima al Alibaba nu a declarat public că face parte a sistemului național de credite, însă calculează cu toate acestea bonitatea unei persoane pe baza unor date precum achizițiile persoanei respective sau prietenii pe care îi are pe rețeaua socială Alipay. În 2015, directorul de tehnologie al Zhima Credit a declarat public că persoanele care cumpără scutece au un „comportament responsabil", în vreme ce cei care joacă jocuri video în exces vor fi penalizați.[116]

Să ne aducem aminte că în introducere am vorbit despre baza de date cloud a poliției din China, construită pentru a monitoriza și urmări persoanele cu tulburări psihice, cele care au criticat guvernul în public și cele care fac parte din minorități etnice. Programul integrat de operațiuni folosește inteligența artificială pentru a detecta abateri de la tiparele normale de comportament, cum ar fi traversarea prin locuri interzise. Sistemul de credite sociale din China notează și ierarhizează cetățenii pe baza comportamentului acestora; sistemele de luare a deciziilor bazate pe inteligență artificială folosesc aceste scoruri pentru a stabili cine se poate împrumuta de la bancă, cine poate călători și chiar școala la care se pot înscrie copiii acestora.

Robin Li, unul dintre fondatorii Baidu, susține că pentru poporul chinez, intimitatea nu este o valoare primordială, așa cum este pentru occidentali. „Poporul chinez este mai deschis sau mai puțin sensibil când vine vorba de viața privată," le-a spus Li celor aflați în public la Forumul de Dezvoltare de la Beijing. „ Dacă au de ales între intimitate și confort, siguranță și eficiență, ei vor prefera în multe situații să facă acest schimb."[117] Sau poate că e vorba mai mult de repercusiuni.

Sunt de părere că sistemul chinez de credite sociale nu urmărește neapărat consolidarea puterii partidului comunist, sau că nu e o modalitate complicată de a dobândi un avantaj strategic față de cei care dezvoltă inteligența artificială în Occident. Acest sistem presupune mai degrabă exercitarea unui control total asupra economiei globale. La începutul anului 2018, președintele Xi a declarat pentru agenția națională de presă Xinhua că „strângând cureaua și din dinți, vom construi ‚două bombe și un satelit'", o referire la un program de înzestrare militară din epoca maoistă. „Asta pentru că am ajuns la performanța maximă a sistemului socialist. Ne-am unit eforturile pentru a realiza lucruri mărețe. Următorul pas este să obținem aceleași rezultate prin știință și tehnologie. Trebuie să renunțăm la iluzii și să ne bazăm doar pe noi înșine."[118]

Xi respinge noțiunea de economie de piață, internet liber și ecosistem divers în care pot coexista idei concurente și complementare. Economia națională a Chinei, puternic centralizată, se izolează de concurență. Ea funcționează pe premisa unor rețele alternative, în care regulile internetului se aplică în funcție de locul unde este situat utilizatorul. Prin intermediul politicilor de

siguranță a datelor, exprimarea liberă este îngrădită, impunându-se o cenzură centralizată în toate domeniile celei de-a treia ere a calculatoarelor: infrastructura internetului, traficul global de date, și aparatura se supun din ce în ce mai mult restricțiilor impuse de Beijing. La o conferință din 2016, Xi a declarat că guvernul intenționează să decidă discreționar cum să protejeze rețelele, aparatura și datele. [119]

Guvernul va exercita acest control puternic oferindu-le partenerilor din Noul Drum al Mătăsii infrastructura și proiecte tehnologice pilot. Tanzania a fost aleasă să participe în etapa pilot de implementare – și, poate că nu e o coincidență, această țară a adoptat acum multe dintre politicile chineze referitoare la date și informatică. Guvernul Tanzaniei beneficiază de asistență tehnică din partea omologilor chinezi, iar un înalt oficial din Tanzania a afirmat că „prietenii noștri chinezi au reușit să blocheze aceste siteuri în propria țară și să le înlocuiască cu platforme echivalente care sunt sigure, populare și constructive."[120] Același lucru se petrece și în alte țări din Africa. Și Vietnamul a adoptat legi draconice în privința securității cibernetice, după model chinezesc. În luna iunie 2018, India avea în vedere să adopte reglementări similare celor chineze referitoare la stocarea locală a datelor și obținerea unor tehnologii naționale de securizare a datelor.[121]

Cum ar fi dacă China ar începe să-și influențeze partenerii din Noul Drum al Mătăsii, iar unul dintre principalele exporturi ar deveni sistemul național de credite sociale? E ușor de imaginat cum regimurile autocrate, cum ar fi Turcia sau Rwanda, ar putea cumpăra această tehnologie de la guvernul chinez. Cum rămâne însă cu alte țări, ca Brazilia sau Austria, care au cedat tentației

populiste și în care în prezent sunt la putere lideri naționaliști? Cum ar fi dacă o agenție guvernamentală din *țara voastră* ar fi inspirată sau forțată să adopte un sistem de credite sociale, care să vă monitorizeze fără acordul vostru explicit? Ați ști că aveți un scor, sau că vă aflați pe o listă de persoane suspecte?

Cum ar fi dacă companiile străine ar fi împărțite în categorii și ar primi fie un tratament preferențial sau li s-ar interzice să facă afaceri în China – sau chiar între ele? Pe măsură ce economia chineză crește, ce s-ar întâmpla dacă această putere și influență s-ar răspândi prin internet, prin aparatele pe care le folosim și chiar prin inteligența artificială însăși?

Cum ar fi dacă China ar construi un sistem de credite sociale pentru persoanele din afara țării, folosind date pe care le extrage pe gratis din internetul liber și rețelele de socializare occidentale? Cum ar fi dacă s-ar folosi de toate datele pe care le lăsați în urma voastră după ce vizitați Marele Zid sau Orașul Interzis? Cum rămâne cu toate atacurile cibernetice de care auzim periodic, în care au loc breșe de securitate care-și au originea în rețele din China?

Există și un alt motiv pentru care ar trebui să ne îngrijoreze planurile Chinei, și aici revenim la incubatorul triburilor inteligenței artificiale: sistemul educațional. China atrage constant profesori și cercetători din centre din Canada și Statele Unite, oferindu-le pachete de relocare atractive. Există deja o penurie de specialiști experimentați în domeniul învățării automate și stocării datelor. Acest exod al creierelor va crea în curând o lipsă de talente în Occident. Aceasta este de departe cea mai inteligentă mutare a Chinei, deoarece va priva Occidentul de capacitatea de a rămâne competitiv în viitor.

Rețeaua de recrutare chineză atrage cercetătorii înapoi în țară ca parte a Planului celor O Mie de Talente. Expansiunea rapidă a grupului BAT a generat o cerere de oameni calificați- majoritatea lor fiind absolvenți în Statele Unite și care lucrează pentru universitățile și companiile din America. Acest plan guvernamental urmărește atragerea specialiștilor și profesorilor, oferindu-le stimulente financiare considerabile (atât de ordin personal cât și pentru proiecte de cercetare) dar și ocazia de a se integra într-un mediu lipsit de constrângerile administrative și juridice întâlnite în Statele Unite. Peste 7.000 de oameni au fost atrași până acum în acest program, primind un bonus la semnare din partea guvernului în sumă de 1 milion de yuan (aproximativ 151.000 de dolari), un buget personal de cercetare de 3-5 milioane de yuan (467.000-778.000 de dolari), subvenții pentru locuință și educație, norme de hrană, pachete de relocare, asistență pentru ca rudele să-și găsească un loc de muncă și chiar bilete de avion pentru ca familia să poată să se reunească.[122] Toți cei care revin în țară își pun – direct sau indirect- talentele în slujba grupului BAT.

Triburile Americii: Mafia G

Dacă inteligența artificială este cursa spațială a Chinei, China are un avans considerabil. În ultimii ani, în timp ce în domeniu s-au făcut descoperiri importante, administrația Trump a redus sumele alocate cercetării, a împrăștiat zvonuri neadevărate despre impactul inteligenței artificiale asupra forței de muncă, a îndepărtat aliații strategici ai Americii și a sâcâit în mod repetat China prin impunerea de tarife vamale.

Ne vom da în curând seama că politicienii noştri nu au o strategie coerentă pentru inteligenţa artificială, sau pentru viitorul nostru pe termen lung. În locul acestei strategii stau oportunismul şi goana după succesul de piaţă. Membrii celor Nouă Titani din Statele Unite ar putea fi companii de succes în mod individual, însă ei nu fac parte a unui demers concertat de a consolida şi centraliza puterea economică şi militară în Statele Unite. Ei nici nu ar putea -sau nu ar trebui – să fie de acord cu un astfel de plan.

Istoria membrilor americani e deja bine cunoscută, însă schimbările semnificative care vor avea loc în relaţia dintre aceste companii, datele voastre şi aparatele pe care le folosiţi sunt mai puţin cunoscute.

Membrii americani-Google, Microsoft, Amazon, Facebook, IBM şi Apple sunt companii inovatoare, inventive care au pus în mare măsură bazele progreselor în inteligenţa artificială. Ei funcţionează ca o *mafie* în sensul cel mai pur (dar nu peiorativ) al cuvântului: constituie o reţea foarte strânsă de oameni cu interese şi provenienţe similare care colaborează în cadrul unui domeniu care ne va afecta viitorul. În momentul de faţă, Google are cea mai mare influenţă în domeniul inteligenţei artificiale, al guvernului şi al vieţii noastre de zi cu zi, aşa că le vom spune companiilor americane Mafia G. Nu e de mirare că China s-a grăbit să le emuleze succesul, sau că li s-a interzis în bună măsură să opereze în ţara respectivă. Aceste companii nu au pornit de la inteligenţa artificială ca element de bază al afacerilor, dar toate aceste şase companii s-au concentrat în ultimii trei ani în mod special asupra aplicaţiilor comerciale ale inteligenţei artificiale, prin intermediul cercetării, al parteneriatelor şi al unor noi produse şi servicii.

În China, guvernul este cel care controlează BAT. În Statele Unite, Mafia G are o putere și o influență semnificativă asupra guvernului, și pentru că în America există o economie de piață, dar și datorită aversiunii culturale puternice împotriva intervențiilor guvernamentale în mediul de afaceri. Există și un alt motiv pentru care Mafia G are o influență atât de mare-membrii ei au fost ignorați de către politicienii de la Washington. În timp ce Xi își consolida poziția și lansa planurile pentru orizontul 2030, asistentul adjunct pentru politici tehnologice al președintelui Trump, Michael Kratsios, declara în fața unui grup de lideri din domeniul tehnologiei reuniți la Casa Albă că pentru Statele Unite este de preferat ca firmele din Silicon Valley să-și decidă singure prioritățile, fără intervenția statului.[123]

Există un dezechilibru al puterii Unite pentru că guvernul Statelor nu a fost capabil să creeze rețelele, bazele de date și infrastructura de care are nevoie pentru a funcționa, astfel că are nevoie de Mafia G. Spre exemplu, există șanse ca Amazon să ofere servicii de cloud computing către guvernul american în valoare de 4,6 miliarde de dolari în 2019- asta în timp ce compania privată spațială a lui Jeff Bezos, Blue Origin, își va oferi probabil sprijinul către NASA și Pentagon în cadrul unor viitoare misiuni spațiale. În Statele Unite, guvernul depinde de Mafia G, și din cauză că suntem o economie de piață care are legi și reglementări menite să protejeze mediul de afaceri, companiile din Silicon Valley exercită o influență considerabilă. Să mă fac înțeleasă: nu neg rolul Mafiei G sau că membrii ei ar fi companii profitabil și de succes. Nici nu cred că faptul că câștigă mulți bani le dă o aură negativă. Mafia G nu ar trebui constrânsă sau reglementată când vine vorba de obținerea profitului, atât timă cât nu încalcă alte legi.

Această oportunitate are însă și costuri. Există o presiune uriașă ca Mafia G să construiască aplicații practice, vandabile ale inteligenței artificiale, și cât mai repede cu putință. În lumea digitală, investitorii s-au obișnuit cu profituri consistente și rapide. Dropbox, o platformă de stocare digitală a ajuns la o valoare de piață de 10 miliarde de dolari după doar 6 ani de la lansare. Firma de investiții din Silicon Valley Sequoia Capital a investit cu o participație de 20% la subscripția inițială a acțiunilor Dropbox, o participație care acum valorează 1,7 miliarde de dolari.[124] În Silicon Valley, companiile noi care ajung la o valoare de peste 1 miliard de dolari sunt cunoscute sub numele de „unicorni", iar cu o valoare de 10 ori mai mare, Dropbox este un „decacorn." În 2018 erau suficienți unicorni și decacorni cât pentru o grădină zoologică în Silicon Valley, iar câțiva dintre ei au parteneriate cu Mafia G, printre care aș aminti SpaceX, Coinbase, Peloton, Credit Karma, Airbnb, Palantir și Uber. Un aflux rapid de bani atrage după sine și așteptări mărite ca produsul sau serviciul să recupereze investiția inițială, fie printr-o adoptare generalizată, o achiziție efectuată de o firmă mai mare sau un renume de piață.

Aveți o relație specială cu Mafia G, chiar dacă nu folosiți produsele lor bine cunoscute. Teoria celor „șase etape de distanță" reprezintă o modalitate matematică de a explica felul în care suntem cu toții înrudiți. Între voi și toți cunoscuții voștri există o distanță de un pas, o distanță de doi pași între voi și cunoscuții cunoscuților și tot așa. Între voi și Mafia G sunt surprinzător de puțini pași, chiar și dacă nu folosiți internetul.

Două treimi dintre adulții din Statele Unite folosesc Facebook,[125] iar majoritatea utilizatorilor accesează rețeaua de sociali-

zare cel puțin o dată pe zi, ceea ce înseamnă că și dacă nu o folosiți, cineva din anturajul vostru o va face. Există cel mult unul sau doi pași între voi și Facebook, chiar și dacă nu ați „apreciat" niciodată o postare, și chiar dacă v-ați șters contul. Aproape jumătate din gospodăriile americane folosesc Amazon Prime, așa că între voi și Amazon există unul până la trei pași distanță.[126] Dacă ați fost la doctor în ultimii zece ani, între voi și Microsoft și IBM e doar un pas distanță. 95% dintre americani au telefoane inteligente,[127] ceea ce înseamnă că între voi și Apple sau Google e doar un pas.

Prin simplul fapt de a fi trăit în ultimele două decenii, ați generat date pentru Mafia G, chiar și dacă nu le folosiți produsele și serviciile. Motivul este că avem un număr uriaș de aparate inteligente care generează date: telefoanele noastre mobile, aparatele GPS, sistemele de comandă vocală inteligente, televizoarele și aparatele de înregistrat filme, camerele de supraveghere, brățările de fitness, senzorii fără fir din grădină și echipamente de fitness interconectate – și asta pentru că o proporție covârșitoare din comunicațiile, cumpărăturile, munca și viața zilnică se petrece pe platformele Mafiei G.

În Statele Unite, firmele pot accesa contra cost aceste date în scopuri comerciale sau pot optimiza funcționalitățile diferitelor sisteme pe care le folosim. Putem cumpăra pe multe siteuri folosind adresa de e-mail și cardul de credit pe care le-am pus pe tavă celor de la Amazon. Ne putem autentifica pe multe portaluri folosind contul de Facebook. Capacitatea de a folosi datele pe care le deține Mafia G pentru alte servicii este legată de datele pe care le generăm- sub forma pozelor, înregistrărilor audio și video, informațiilor biometrice, ale traficului de internet și așa mai de-

parte. Toate datele noastre sunt stocate în „cloud", un termen în vogă care descrie programele și serviciile care sunt stocate pe internet și nu pe propriile voastre aparate. Poate că nu e de mirare că principalii furnizori de servicii cloud sunt Google, Amazon, Microsoft și IBM.

Accesați serviciile cloud direct (când creați documente Google partajate) sau indirect (atunci când telefonul vostru creează automat o copie de siguranță a fotografiilor pe serverele companiilor). Dacă aveți un iPhone sau un iPad, folosiți serviciul cloud al Apple. Dacă ați accesat siteul healthcare.gov în Statele Unite, ați accesat serviciul cloud al Amazon. Dacă copilul vostru a avut o petrecere de zi de naștere Build-A-Bear la mall, ea a fost organizată prin serviciul cloud al Microsoft. În ultimul deceniu, serviciile cloud au luat amploare- una atât de mare încât nici nu ni se mai par remarcabile sau interesante sau stimulante din punct de vedere tehnic. Ele există pur și simplu, la fel ca electricitatea sau apa curentă. Nu ne gândim mai deloc la ele decât când nu le mai putem accesa.

Cu toții generăm date și folosim serviciile cloud, având o încredere oarbă în triburile inteligenței artificiale și sistemele comerciale create de acestea. În Statele Unite, datele noastre spun mult mai multe despre noi decât codul de asigurări sociale pe care am fost învățați să-l păzim cu sfințenie. Pe baza numărului de asigurări sociale, o persoană poate să deschidă un cont bancar sau să solicite un credit pentru cumpărarea unei mașini. Cu datele stocate în cloud, Mafia G ar putea teoretic să știe dacă ascundeți o sarcină, dacă angajații cred că sunteți un patron incompetent sau dacă aveți o boală terminală-iar inteligența artificială a Mafiei G

știe probabil toate aceste lucruri cu mult înainte ca voi să le aflați. Perspectiva divină pe care o are Mafia G asupra vieților noastre nu e neapărat un lucru rău. Există multe feluri prin care minarea datelor noastre personale pentru a afla informații noi ar putea avea ca rezultat o viață mai fericită și mai sănătoasă pentru toți.

Oricât de puternice ar fi serviciile cloud și inteligența artificială a Mafiei G, ele au încă anumite limitări: aparatura. Arhitectura inteligenței artificiale din zilele noastre a fost suficient de bună pentru a construi produse cu inteligență artificială limitată, cum ar fi filtrul de mesaje nedorite al Gmail, sau serviciul Apple de convertire a mesageriei vocale în text. Această arhitectură va trebui însă adaptată la inteligența artificială generală, o etapă mai dificilă care se întrevede în acest moment la orizont. Pentru a putea trece la următorul nivel va fi nevoie de o aparatură special creată pentru inteligența artificială.

Motivul este legat de un element amintit anterior de John von Neumann, informaticianul care a formulat teoria care stă la bazele arhitecturii calculatoarelor moderne. Trebuie să vă aduceți aminte că pe vremea lui von Neumann, calculatoarele primeau programe și date separate pentru a fi procesate- în această arhitectură programele și datele erau stocate simultan în memoria aparatului. Această arhitectură este folosită și în zilele noastre pentru laptopuri și computere de birou, iar datele se mișcă între procesor și memorie. Dacă nu aveți suficientă putere de calcul, aparatul se va încinge sau veți primi un mesaj de eroare sau calculatorul se va bloca. Problema este cunoscută sub numele de „gâtul de sticlă von Neumann." Indiferent cât de repede ar putea procesorul să prelucreze datele, memoria programului și cea a datelor provoacă acest

gât de sticlă care limitează rata de transfer. Mai toate calculatoarele din ziua de azi sunt bazate pe această arhitectură von Neumann, iar problema constă în faptul că procesoarele din zilele noastre nu pot rula programele mai rapid decât pot prelua instrucțiunile și datele din memorie.

Gâtul de sticlă constituie o problemă mare a inteligenței artificiale. Dacă vorbiți cu sistemele Alexa sau Google Home, vocea voastră este înregistrată, convertită în date și apoi transmisă serverului din cloud pentru o reacție – dată fiind distanța fizică între voi și diferitele locuri unde se găsesc serverele implicate, e uluitor că Alexa poate să reacționeze în doar o secundă sau două. Pe măsură ce tot mai multe aparate capătă o componentă legată de inteligența artificială – sub forma telefoanelor inteligente cu senzori biometrici, camerelor de supraveghere care ne pot recunoaște fața, mașinilor care se conduc singure, sau a roboților preciși care pot doza medicamente- o întârziere de procesare de doar o secundă sau două poate provoca un rezultat catastrofal. O mașină autonomă nu poate contacta serverul cloud pentru fiecare acțiune, pentru că sunt mult prea mulți senzori care ar trebui să transmită datele către serverul cloud pentru a fi procesate.

Singura soluție o reprezintă deplasarea puterii de calcul mai aproape de sursa datelor, ceea ce va conduce la reducerea timpului de răspuns și a încărcării rețelelor. Această nouă arhitectură se numește „edge computing" și reprezintă evoluția inevitabilă a aparatelor dotate cu inteligență artificială și a arhitecturii sistemelor. Pentru ca inteligența artificială să progreseze la următoarea etapă, aparatura trebuie să țină pasul cu ea. În loc să ne întâlnim cu Mafia G pe serverul cloud, unde încă mai avem o capacitate de a modi-

fica permisiuni și setări, vom fi nevoiți să-i invităm să controleze *toate* aparatele pe care le folosim. Aceasta înseamnă că în următorul deceniu, restul ecosistemului inteligenței artificiale va funcționa doar pe baza câtorva sisteme ale Mafiei G. Toate companiile nou înființate și firmele marginale- ca să nu mai vorbim de oamenii ca noi- vor trebui să accepte o nouă ordine și să jure credință unei mâini de furnizori de servicii care vor furniza sistemele de operare pentru viața de zi cu zi. Odată ce datele, aparatele, mașinile și serviciile vor fi interconectate, veți deveni captivi. Pe măsură ce cumpărați mai multe bunuri – telefoane mobile, frigidere cu conexiune la internet, căști inteligent – veți constata că Mafia G vă oferă sistemul de operare pentru viața zilnică. Omenirea primește o ofertă pe care nu o poate refuza.

Calculele legate de învățarea profundă necesită aparatură specializată pentru că au nevoie de multă putere de calcul. Deoarece acestea preferă optimizarea în favoarea preciziei și se compun în esență din operațiuni complexe de algebră liniară, este de înțeles că o nouă arhitectură a rețelelor neurale ar provoca o eficiență mărită, dar, și mai important, o accelerare a procesului de design și lansare pe piață. Cu cât echipele de cercetare pot construi și testa modele reale, cu atât se vor apropia mai mult de aplicații practice ale inteligenței artificiale. Spre exemplu, în prezent e nevoie de săptămâni sau de luni pentru a pregăti un model complex de vizualizare digitală- iar rezultatul final ar putea doar să dovedească nevoia unor ajustări suplimentare, ceea ce înseamnă că munca trebuie reluată de la capăt. O aparatură mai bună ar presupune ca aceste modele să poată fi pregătite în câteva ore, sau câteva minute, ceea ce ar conduce la noi descoperiri săptămânale, sau chiar zilnice.

Acesta este motivul pentru care Google și-a creat propriile circuite, denumite Unități de Procesare prin Tensori (UPT). Aceste circuite pot funcționa cu structura de învățare profundă prin inteligență artificială, TensorFlow. În luna iunie 2018, TensorFlow era principala platformă de învățare automată de pe GitHub, care este cea mai mare platformă de pe internet din lume în care programatorii își depozitează softul. TensorFlow a fost descărcată de peste 10 milioane de ori de către programatori din 180 de țări, la momentul redactării acestei cărți existând 24.500 de copii active ale platformei. [128] Google a completat platforma cu produse suplimentare, cum ar fi TF-GAN (o bibliotecă de module pentru rețele neurale opuse)[c] și programul TensorFlow de detectare a obiectelor (care îi ajută pe programatori să creeze modele mai exacte de învățare pentru vizualizarea digitală). Circuitele UPT sunt deja folosite în serverele Google- ele stau la baza modelelor de învățare profundă care se activează cu fiecare căutare de pe Google.

Google avea motive întemeiate să încerce să achiziționeze GitHub, o platformă cu 28 de milioane de programatori din lumea întreagă și cu legături puternice cu ceilalți membri ai celor Nouă Titani. În luna iunie 2018, Google a fost depășită de – cine credeți? – Microsoft.[129]

Facebook a colaborat cu Intel pentru dezvoltarea unui chip de inteligență artificială pentru propria activitate de cercetare de care compania avea nevoie pentru a accelera procesul de experimentare. Apple și-a creat propriul circuit „cu motor neural" pe

---

c Rețele neurale la care învățarea se produce din interacțiunea între rețele opuse una celeilalte.(n.t)

care îl folosește în cadrul telefoanelor iPhone X, iar Microsoft a dezvoltat circuite de inteligență artificială pentru casca de realitate virtuală HoloLens și platforma de cloud computing Azure. Membrii BAT își construiesc de asemenea propriile circuite: Alibaba a început un proces intens de recrutare în Silicon Valley, căutând „arhitecți pentru circuite de inteligență artificială"[130], iar în 2018 și-a lansat propriile circuite – Ali-NPU – pe care oricine le poate folosi prin intermediul serviciului public cloud.

Anticipând faptul că în viitor cerințele de performanță vor crește, IBM a dezvoltat acum câțiva ani circuitul neuromorfologic TrueNorth, lucrând deja la un nou tip de aparate care ar putea spori eficiența rețelelor neurale de peste 100 de ori. Ca să înțelegeți despre ce e vorba, e ca o diferență de la un abac la mașina de teleportare din Star Trek. Această nouă generație de circuite folosește două tipuri de sinapse, una pentru memoria pe termen lung, și alta pentru calcule pe termen scurt.

Vorbim despre un echivalent al zilelor noastre al întrebării „ești genul de om care folosește un PC sau un Mac?" dus la extrem. Majoritatea acestor circuite funcționează pe baza unor structuri pe care cei Nouă Titani le definesc a fi libere de drepturi de autor- ceea ce înseamnă că dezvoltatorii pot să acceseze, să folosească și să îmbunătățească gratuit structurile. Circuitele fizice sunt însă protejate, iar serviciile sunt bazate pe abonamente. În practică, aceasta înseamnă că odată ce o aplicație este construită pentru o anumită structură, ea va fi deosebit de dificil de adaptat pentru o altă structură. Acesta este modul prin care triburile inteligenței artificiale recrutează noi aderenți-iar ritualul de inițiere este sărutarea inelului unei structuri a Mafiei G.

Căutând să comercializeze inteligența artificială, Mafia G recrutează programatori într-o manieră creativă. În luna mai 2018, Google și platforma de cursuri online Coursera au lansat o nouă specializare dedicată învățării automate. Pentru a participa, trebuie însă să folosiți TensorFlow. Cursul cu cinci părți, la finalul căruia obțineți un certificat de absolvire, este descris ca o modalitate accesibilă oricui de a învăța despre învățarea automată și rețelele neurale. Cursanții au nevoie de date și structuri din lumea reală, așa că se bazează pe structura pusă la dispoziție de Google.

Aparatura face parte din strategia pentru inteligență artificială a Mafiei G, care are legături cu guvernul, diferite de cele despre care am vorbit în cazul Chinei, dar care ar trebui să vă provoace îngrijorări similare, chiar și dacă nu sunteți cetățean al Statelor Unite.

Motivul este că în Statele Unite, inteligența artificială este pusă în slujba a trei stăpâni: guvernul, lumea finanțelor de pe Wall Street și companiile din Silicon Valley. Oamenii care redactează politicile guvernamentale sau reglementările sunt tehnocrați de carieră, care au tendința să-și păstreze zeci de ani slujbele. Cei care dețin însă controlul – președintele sau șefii marilor agenții guvernamentale (Federal Communications Commission, Departamentul de Justiție și alții) se schimbă din funcție o dată la câțiva ani. În domeniul inteligenței artificiale nu există o direcție clară la nivel național.

Doar recent a început să fie luată în calcul mai clar poziția Chinei și planurile acesteia în domeniul inteligenței artificiale, în principal din cauza faptului că președintele Xi a dat publicității planul strategic pe termen lung asupra inteligenței artificiale și folosirii datelor. În Statele Unite există o autoritate denumită Comi-

tetul pentru Investiții străine. Acesta este o comisie compusă din reprezentanți ai puterii și opoziției condus de secretarul Trezoreriei și compus din membri ai Departamentelor Trezorerie, Justiției, Energiei, Apărării, Comerțului, de Stat și al Siguranței Naționale. Misiunea lor este de a analiza și investiga tranzacțiile comerciale care ar putea periclita siguranța națională. Acest comitet a blocat achiziția companiei Qualcomm, un fabricant de circuite din San Diego, de către firma Broadcom din Singapore. Același comitet a respins și achiziția companiei MoneyGram din Dallas de către compania de plăți electronice Ant Financial, deținută de Alibaba. La momentul scrierii cărții, comitetul nu acordă o atenție specială inteligenței artificiale, deși s-a propus ca puterile sale să fie sporite pentru a preîntâmpina expansiunea Chinei prin cumpărarea companiilor americane.

Între timp, în Silicon Valley a devenit o practică curentă ca angajații să treacă de la o compania la alta, în timp ce liderii tribului își împart timpul între Mafia G și au o stabilitate mai mare a slujbei. Inteligența artificială continuă să progreseze în timp ce mantra tribului -*construiește acum, cere-ți scuze mai târziu*- devine din ce în ce mai puternică. Ani de zile Google a scanat și catalogat cărți protejate de drepturi de autor fără să ceară în prealabil permisiunea autorilor- compania a fost acționată în justiție de edituri și autori. Google a fotografiat casele și cartierele noastre și le-a publicat pe Google Maps, fără să ne ceară permisiunea. (Oamenii sunt evitați, pe cât posibil, iar chipurile lor sunt voalate). Apple încetinește deliberat performanța telefoanelor mai vechi când lansează modele noi, apoi își cere scuze. După scandalul Cambridge Analytica, directorul general al Facebook Mark Zuckerberg a pu-

blicat o dezmințire pe propriul zid de Facebook, scriind „Celor pe care i-am dezamăgit, le cer iertare și voi încerca să fiu mai bun. Cer iertare pentru felul în care munca mea a contribuit la divizarea oamenilor în loc să ne aducă împreună."

Prin urmare, Mafia G are tendința de a acționa cu rapiditate în salturi evolutive până când se petrece ceva rău, apoi se implică guvernul. Politicile Facebook de securitate a datelor nu au atras atenția guvernului decât în momentul în care un fost angajat Cambridge Analytica a rupt tăcerea, explicând ușurința cu care au fost adunate și vândute datele noastre. În 2016, ca urmare a unui atac armat petrecut la San Bernardino, California, guvernul federal a încercat să oblige Apple să ofere autorităților o cheie de decriptare a iPhone-ului atacatorului. Agențiile guvernamentale și oamenii legii susțineau că decodarea telefonului și transmiterea datelor private era în interesul publicului, în timp ce apărătorii dreptului la intimitate susțineau că ar viola drepturile civile. În cele din urmă, autoritățile au reușit să spargă telefonul fără ajutorul Apple, așa că nu am apucat să vedem care tabără a avut dreptate. În Statele Unite intimitatea este la mare preț, dar nu avem legi clare, care să lămurească situația datelor personale în secolul 21.

În vara anului 2018, personalul de la cabinetul senatorului Mark Warner (D-VA) a lansat în public o serie de propuneri menite să înfrâneze avântul giganților din domeniul tehnologiei. Printre acestea se numără crearea unei legislații asemănătoarea cu reglementările GDPR din Europa sau propunerea ca platformele de internet să fie considerate persoane de încredere care trebuie să urmeze un anumit cod de conduită, asemănător cu al firmelor de avocatură.[131] La doar câteva luni după aceea, directorul general

al Apple, Tim Cook, a postat pe Twitter câteva gânduri despre viitorul intimității, coloșii din tehnologie și America. Pe 24 octombrie el a scris că protecția intimității utilizatorilor ar trebui să reprezinte principala prioritate a companiilor. „Companiile trebuie să recunoască faptul că datele aparțin utilizatorilor și ar trebui să înlesnim capacitatea oamenilor de a a obține o copie a datelor lor personale, precum și de a le corecta sau de a le șterge", scria el, continuând, „Oricine are dreptul ca datele proprii să fie în siguranță."[132] Simțind că în Statele Unite reglementările în domeniul devin plauzibile, Apple și-a promovat serviciile de protecție a datelor precum setările de intimitate incluse în sistemele de operare pentru terminale mobile și computere.

Suntem de acord să fim supravegheați constant în schimbul serviciilor de care beneficiem. Acest troc permite Mafiei G să facă profit, pentru a-și putea îmbunătăți și extinde paleta de servicii pe care ni le oferă, fie în calitate de consumatori individuali, fie ca firme, universități, organizații neguvernamentale sau agenții ale statului. Avem de-a face cu un model de afaceri bazat pe capitalismul de supraveghere. Și, dacă e să fim sinceri, e un sistem pe care în Statele Unite l-am acceptat, altminteri am fi încetat de mult să mai folosim servicii ca Gmail, Microsoft Outlook sau Facebook. Pentru a putea funcționa cum trebuie, ei au nevoie de datele noastre, care sunt extrase, procesate și ambalate. Presupun că folosiți cel puțin unul dintre produsele sau serviciile oferite de către Mafia G. Eu folosesc zeci de astfel de produse și sunt perfect conștientă care este prețul lor real.

Ceea ce vreau să spun este că în curând nu vom încredința Mafiei G doar datele noastre. Pe măsură ce trecem de la o

inteligență artificială limitată la una mai vastă, capabilă să ia decizii complexe, vom invita aceste companii direct în frigiderele, dulapurile și mașinile noastre, și în ochelarii, brățările și căștile interconectate pe care le vom purta în curând. Astfel, Mafia G va putea să automatizeze muncile repetitive, să ne ajute să luăm decizii și să ne ajute să gândim mai ușor. Între noi și Mafia G nu va mai exista nici un pas de distanță. Politicienii nu vor mai putea invoca nicio autoritate reală odată ce întreaga noastră existență va fi strâns legată de aceste companii. La ce am putea însă renunța, în schimb?

Cei nouă titani – grupul BAT (Baidu, Alibaba și Tencent) și Mafia G (Google, Microsoft, Amazon, Facebook, IBM și Apple) – dezvoltă instrumentele și mediul care vor sta la baza viitorului inteligenței artificiale. Aceste companii fac parte din tribul inteligenței artificiale, format în universități în care se induc idealuri și obiective comune, care se rigidizează și mai mult odată cu intrarea absolvenților pe piața muncii. Domeniul inteligenței artificiale nu este unul static. Pe măsură ce inteligența artificială îngustă evoluează către o inteligență artificială generală, cei Nouă Titani dezvoltă noi tipuri de sisteme de calcul și recrutează programatori care devin prizonieri ai structurilor lor.

Modelul consumist al inteligenței artificiale sub forma întâlnită în Statele Unite nu este în mod obligatoriu rău, așa cum nici modelul chinez, bazat pe intervenția centralizată a statului nu este. Inteligența artificială nu este neapărat dăunătoare societății. Cu toate acestea, Mafia G se compune din companii care urmăresc profiturile, companii listate la bursă care trebuie să dea socoteală

acționarilor, indiferent de intențiile altruiste ale conducătorilor și angajaților lor. În China, grupul BAT se supune guvernului chinez, care a stabilit deja prioritățile cetățenilor. Ceea ce aș vrea să știu – și ceea ce ar trebui să vă intereseze și pe voi- este ce e mai bine pentru toată omenirea? Pe măsură ce inteligența artificială evoluează, cum se vor reflecta deciziile pe care le luăm astăzi în deciziile pe care mașinile le vor lua în numele nostru în viitor?

CAPITOLUL TREI

O MIE DE TĂIETURI CU HÂRTIA: CONSECINȚELE NEINTENȚIONATE ALE INTELIGENȚEI ARTIFICIALE

„Mai întâi ne formăm obiceiurile, apoi obiceiurile ne formează pe noi."
—JOHN DRYDEN

„Tu ești creatorul meu, dar eu sunt stăpânul tău."
—MONSTRUL LUI FRANKENSTEIN (DE MARY SHELLEY)

Contrar tuturor poveștilor apocaliptice pe care le-ați citit și văzut în care inteligența artificială se răscoală dintr-o dată și se decide să elimine omenirea, nu va exista un singur eveniment în care tehnologia să dea greș. Ceea ce urmează să trăim cu toții e mai degrabă ca o serie treptată de tăieturi cu hârtia. Dacă vă tăiați o dată la deget, e enervant, dar vă puteți vedea în continuare de treabă. Dacă tot corpul e plin de mii de mici tăieturi, nu veți muri, dar viața va deveni un chin. Micile amănunte ale vieții-încălțatul, a mânca un taco, a dansa la nunta unui văr-nu ar mai fi alegeri. Veți deveni nevoiți să învățați să trăiți un alt fel de viață. O viață plină de restricții. O viață cu consecințe dureroase.

Știm deja că etica și diversitatea nu sunt obligatorii în universitățile unde se formează triburile inteligenței artificiale și nici în cadrul celor Nouă Titani, unde lucrează mai apoi triburile inteligenței artificiale. Știm că Mafia G funcționează prin accelerarea acelor proiecte și studii care au șanse de a fi vândute pe piață, iar grupul BAT urmărește îndeplinirea unui plan centralizat al guvernului chinez. Devine evident că poate nimeni – nu o agenție mondială de reglementare (asemănătoare Agenției Internaționale pentru Energia Atomică), sau un grup de universități sau măcar un grup de cercetători- nu se gândește la întrebările dificile legate de ecartul care se creează, la conflictul dintre valorile oamenilor și valoarea economică a planurilor chineze de supremație în domeniul inteligenței artificiale, sau obiectivele mercantile ale companiilor din Silicon Valley. Echilibrul dintre valorile umane și obiectivele economice nu a constituit o prioritate în trecut deoarece cei Nouă Titani au adus prosperitate, ne oferă servicii și produse bune pe care le folosim cu toții și ne fac să ne simțim stăpâni pe domeniul nostru digital. Nu am cerut răspunsuri legate de valori pentru că momentan viețile noastre ni se par mai bune datorită celor Nouă Titani.

Am fost însă deja tăiați cu hârtia de convingerile și motivațiile creatorilor inteligenței artificiale. Cei Nou Titani nu se ocupă doar cu aparatura și programele. Ei construiesc mașini gânditoare care reflectă valorile omenirii. Ecartul prezent dintre triburile inteligenței artificiale și oamenii obișnuiți produce deja rezultate îngrijorătoare.

Algoritmul valorilor

V-ați întrebat vreodată de ce sistemul inteligenței artificiale nu e mai transparent? V-ați întrebat ce fel de date sunt folosite- inclusiv datele voastre personale- pentru a ajuta inteligența artificială să învețe? În ce situații este învățată inteligența artificială să facă excepții? Cum reușesc să împace creatorii mercantilizarea inteligenței artificiale cu nevoile omenești primare, cum ar fi intimitatea, siguranța, sentimentul de apartenență, respectul de sine și auto-actualizarea? Care sunt imperativele morale ale tribului inteligenței artificiale? Ce cred ei că e bine și rău? Reușesc ei să învețe inteligența artificială ce este empatia? (La drept vorbind, e util sau merită să încercăm să învățăm inteligența artificială ce este empatia omenească?)

Fiecare dintre cei Nouă Titani are un set formal de valori, însă aceste coduri nu răspund la aceste întrebări. În schimb, aceste valori declarate sunt convingeri profunde care unesc, inspiră și motivează angajații și acționarii. Valorile unei companii acționează ca un algoritm – un set de reguli și instrucțiuni care influențează cultura organizațională, stilul de conducere și joacă un rol important în toate deciziile luate, de la ședințele conducerii până la liniile individuale ale unui program. Absența unor anumite valori afirmate este remarcabilă și pentru că dacă ele exprimate explicit, devin greu de sesizat și ușor de uitat.

La început, Google funcționa pe baza unei valori simple, fundamentale: „nu fii rău."[133] În scrisoarea de intenție de la subscripția publică de acțiuni în 2004, fondatorii Sergey Brin și Larry page scriau: Eric [Schmidt], Sergey și cu mine intenționăm ca Google

să fie un alt fel de companie, care va pune în practică valorile pe care le-a acumulat când era o companie deținută privată și în viitorul ei de companie listată la bursă... Ne va interesa un rezultat pe termen lung, fără să urmărim în mod special veniturile trimestriale. Vom sprijini anumite proiecte riscante dar cu potențial mare și vom gestiona portofoliul de proiecte... Ne vom ridica la înălțimea principiului ‚nu fii rău' menținând încrederea utilizatorilor."[134]

„Principiile de conducere" ale Amazon se regăsesc în structura de conducere, iar aceste valori se sprijină pe încredere, metrici, viteză și rezultate. Printre principiile publicate se numără:

> „Liderii pornesc de la utilizator și se mulează pe el. Ei vor lucra intens pentru a dobândi și a păstra încrederea clientului..."
> „Liderii au constant standarde înalte" pe care cei din exterior le-ar putea considera „nerezonabil de înalte."
> „Multe decizii și acțiuni sunt reversibile și nu necesită o analiză amănunțită. Prețuim asumarea calculată a riscurilor."
> „Fă mai mult cu mai puțin. Nu veți câștiga puncte în plus dacă sporiți artificial numărul de angajați, dimensiunea bugetului sau cheltuielile fixe."[135]

Facebook enumeră „5 valori fundamentale," respectiv „a fi îndrăzneț", „a te concentra asupra impactului", „a acționa rapid", „a fi transparent" față de activitatea companiei și „a aduce valoare" pentru utilizatori.[136] Pe de altă parte, „filosofia de management" a Tencent pune accent pe „instruirea și încurajarea angajaților pentru a obține reușite" pe baza „unei atitudini de încredere și respect" și luarea deciziilor pe baza unei formule intitulate „Integri-

tate+Proactivitate+Colaborare+Inovare."[137] Alibaba pune mai presus de orice „o atenție constantă acordată nevoilor clienților noștri", precum și munca în echipă și integritatea.[138]

Dacă aș desena o diagramă Venn a tuturor valorilor și principiilor de funcționare ale celor Nouă Titani, am observa câteva zone în care ele se suprapun. Toate aceste companii pretind angajaților și echipelor să se perfecționeze constant, să construiască produse și servicii de care clienții să nu se poată lipsi și să îndeplinească așteptările acționarilor. Mai presus de orice, ele apreciază încrederea. Aceste valori nu sunt deosebite-ele par a fi valorile majorității companiilor americane.

Pentru că inteligența artificială are potențialul de a avea un impact major asupra omenirii, valorile celor Nouă Titani ar trebui precizate explicit- iar noi ar trebui să le pretindem un standard mai înalt decât celorlalte companii.

Ceea ce lipsește este o formulare neechivocă prin care să se afirme faptul că omenirea reprezintă punctul central al evoluției inteligenței artificiale și că toate demersurile din acest domeniu ar trebui să se axeze pe îmbunătățirea condiție umane. Acest lucru ar trebui afirmat explicit- iar această formulare ar trebui să se regăsească și în alte documente ale companiilor, în ședințele de conducere, în cadrul echipelor care lucrează în inteligența artificială și în cadrul discuțiilor de marketing și vânzări. Spre exemplu, ar trebui ca dincolo de inovație și eficiență să se discute și despre accesibilitate. Milioane de oameni au handicapuri: au probleme de vorbire, auz, văz, tactile sau de gândire. Sau să se discute despre valorile economice, despre puterea platformelor de a crește și de a crea prosperitate fără să excludă anumite persoane sau grupuri.

Sau despre valorile sociale, cum sunt integritatea, diversitatea, toleranța și curiozitatea.

În timp ce scriam această carte, directorul general al Google, Sundar Pichai, a anunțat că Google a redactat un nou set de principii fundamentale referitoare la demersurile companiei în domeniul inteligenței artificiale. Totuși, aceste principii nu merg atât de departe încât să definească omenirea ca fiind punctul de referință pentru demersurile viitoare ale Google în acest domeniu. Comunicarea acestor principii nu constituie o realiniere strategică a valorilor fundamentale ale companiei, ci o reacție datorată scandalului Project Maven-dar și unui incident petrecut la începutul anului. Un grup de ingineri cu experiență a descoperit că un proiect la care lucrau – o metodă de siguranță pentru serviciile de tip cloud – era menită să ajute compania să câștige contracte cu armata. Amazon și Microsoft aveau deja un grad de încredere ridicată pentru un serviciu de tip cloud dedicat guvernului, deținând astfel autorizația de stocare a unor date confidențiale. Google voia să câștige contracte bănoase cu Departamentul Apărării, iar în momentul în care inginerii au aflat, s-au răzvrătit. Ca urmare a aceste răzvrătiri, 5% dintre angajații Google au condamnat public proiectul Maven.[139]

Acesta a fost începutul unei serii de proteste care a început în 2018, când o parte a tribului și-a dat seama că munca lor era pusă în folosul unei cauze pe care nu o sprijineau, așa că a solicitat o schimbare. Ei au presupus că valorile lor personale erau și cele ale companiei-și atunci când și-au dat seama că lucrurile nu stau așa, au protestat. Acest incident demonstrează problemele spinoase care pot apărea atunci când membrii Mafiei G nu respectă standarde mai înalte decât companiile care oferă produse mai puțin spectaculoase.

Nu e deci de mirare că o bună parte din principiile Google legate de inteligența artificială vizează armele și contractele cu armata: Google nu va crea tehnologii care să poată fi folosite militar și al căror principal obiectiv este rănirea altor oameni, Google nu va crea sisteme de inteligență artificiale care să contravină principiilor general acceptate de drept internațional, și așa mai departe. „Vrem să fie clar că deși nu elaborăm sisteme de inteligență artificială care să fie folosite pentru a crea arme, vom continua să colaborăm cu guvernul și armata", susținea documentul.[140]

Spre meritul său, Google susține că principiile sunt menite să reprezinte standarde concrete, nu doar noțiuni teoretice – iar documentul adresează problema prejudecăților din seturile de date. În document nu se amintește însă nicăieri problema transparenței decizionale a inteligenței artificiale, sau a căror seturi de date sunt folosite. Nu se vorbește de faptul că în domeniul inteligenței artificiale lucrează pentru Google triburi omogene. Niciunul dintre aceste standarde concrete nu pune interesele omenirii mai presus de cele ale bursei.

Problema este una de transparență. Dacă guvernul Statelor Unite nu este capabil să construiască sistemele de care avem nevoie pentru a ne păstra siguranța națională, e de așteptat ca el să apeleze la o companie care să poată face acest lucru – și lucrurile au funcționat în acest mod încă de la Primul Război Mondial. Am uitat prea ușor că pacea se câștigă constant și că doar o armată bine pregătită ne garantează siguranța și securitatea națională. Ministerul Apărării nu este o instituție tiranică, care să-și dorească super-arme bazate pe inteligența artificială pentru a rade de pe fața pământului așezări din afara țării. Armata Statelor Unite are misiuni dincolo de uciderea oamenilor și aruncarea în aer a obiec-

tivelor inamice. Dacă cei care lucrează în cadrul Mafiei G nu au înțeles prea bine acest aspect, asta se datorează și faptului că prea puțini oameni au reușit să facă o punte între guvern și Silicon Valley.

Ar trebui să reflectăm cu toții la faptul că cei Nouă Titani construiesc sisteme care se bazează pe oameni, dar valorile care exprimă dorința noastră de a îmbunătăți calitatea vieții oamenilor nu sunt precizate în mod explicit. Dacă valorile tehnologice, economice și sociale nu fac parte din carta de valori a unei companii, este puțin probabil ca acea companie să urmărească binele omenirii în procesul de cercetare, dezvoltare și lansare pe piață. Acest ecart al valorilor nu este întotdeauna evident într-o organizație, iar acest fapt implică un risc semnificativ atât pentru Mafia G cât și pentru BAT, deoarece el îndepărtează angajații de potențialele rezultate negative ale activității lor. Când indivizii și echipele nu sunt conștienți de acest ecart al valorilor, ei nu vor lua măsuri în privința unor probleme esențiale când produsele sunt construite, testate, promovate, lansate și vândute. Nu înseamnă că persoanele care lucrează în domeniul inteligenței artificiale ar fi indiferente-dar înseamnă că ele nu țin cont de valorile de bază ale omenirii.

Și așa ajungem să ne tăiem cu hârtia.

Legea lui Conway

Informatica, la fel ca orice alt domeniu al tehnologiei, sau orice altă activitate, reflectă perspectiva asupra lumii și experiențele echipei care se ocupă de inovație. Vedem aceste lucruri și în afara tehnologiei. Permiteți-mi să mă îndepărtez de inteligența artificială preț de o clipă și să vă prezint două exemple aparent dis-

tincte în care un mic trib de persoane poate ajunge să dobândească o putere extraordinară asupra unei populații.

Dacă aveți părul drept-gros, aspru, fin, lung, scurt, subțire (sau chiar rar) – veți avea o cu totul altă experiență la coafor față de mine. Fie că mergeți la o frizerie sau la un salon cu pretenții, mai întâi vi se spală părul într-o mică chiuvetă, în care cineva trece cu degetele prin părul vostru. Frizerul sau stilistul folosește apoi un pieptene des ca să vă întindă părul în linii drepte. Dacă aveți mult păr, stilistul ar putea folosi o perie și un uscător de păr, trăgând din nou de fiecare șuviță până când are forma dorită – plină de volum și elastic, sau plată și netedă. Dacă aveți părul mai scurt, peria e mai mică și timpul de uscare mai scurt, dar procesul e în linii mari același.

Părul meu e foarte creț și foarte des. Se încurcă ușor și reacționează imprevizibil la factorii de mediu. În funcție de umiditate, de nivelul de hidratare și de produsele cosmetice pe care le-am folosit, părul meu stă bine sau poate fi un dezastru încâlcit. La un salon obișnuit, chiar și acelea unde nu am avut probleme, chiuveta îmi dă de furcă. Persoana care-mi spală părul are de regulă nevoie de mult mai mult spațiu decât permite chiuveta, și uneori părul se încâlcește în jurul furtunului, moment în care devine dureros de descâlcit. Singura modalitate prin care un piepten obișnuit poate trece prin părul meu e dacă părul e ud și acoperit de un lichid alunecos, cum ar fi un balsam mai gros. (De perie nici nu poate fi vorba). Puterea unui uscător obișnuit de păr mi-ar face părul să se înnoade. Unele saloane au un capăt special care împrăștie aerul – arată ca un bol de plastic cu niște mici protuberanțe ca niște ardei iuți- dar pentru a-l putea folosi eficient trebuie

să-mi las capul în jos, iar stilistul trebuie să se aplece ca să poată poziționa corect uscătorul.

Aproximativ 15% dintre albi au părul creț. Dacă adunăm și populația negrilor, rezultă cam 79 de milioane de oameni, aproximativ un sfert din populația Statelor Unite, oameni care au probleme la frizer pentru că, putem presupune, instrumentele și mediul au fost create de oameni cu părul drept, care nu au ținut cont de valori sociale, cum ar fi empatia și diversitatea, în cadrul companiilor lor.[141]

Acesta e un exemplu destul de inofensiv. Gândiți-vă acum la o situație în care mizele sunt mult mai mari decât o simplă coafură. În luna aprilie 2017, agenții de îmbarcare de la un zbor supra-aglomerat al United Airlines de la aeroportul internațional Chicago O'Hare au făcut un anunț la difuzor, cerând pasagerilor să cedeze locurile lor angajaților companiei în schimbul unei compensații de 400 de dolari și o cameră gratuită la un hotel din zonă. Nimeni nu a acceptat oferta. Angajații au crescut apoi oferta la 800 de dolari plus camera de hotel, dar, din nou, nimeni nu a acceptat. Între timp, pasagerii cu prioritate la îmbarcare începuseră să se îmbarce în avion, inclusiv cei de la clasa I.

Un algoritm și un sistem automat a ales să elimine patru oameni, printre care se număra doctorul David Dao și soția lui, de asemenea medic. El le-a spus personalului liniei aeriene că are pacienți pe care trebuie să-i consulte a doua zi. În timp ce ceilalți pasageri au acceptat să fie îndepărtați, Dao a refuzat să se miște. Autoritățile l-au amenințat pe Dao cu închisoarea dacă refuză să plece. Ați văzut cu siguranță ce a urmat pentru că videoclipul incidentului a devenit viral pe Facebook, YouTube și Twitter și a fost

retransmis apoi zile la rând pe canalele de ştiri din toată lumea. Poliţiştii l-au apucat pe Dao de braţe şi l-au scos din scaun cu forţa, lovindu-l de o cotieră, distrugându-i ochelarii şi provocându-i o tăietură în zona gurii. Cu faţa plină de sânge, Dao s-a oprit dintr-o dată din urlete în timp ce poliţia îl târa pe coridorul avionului United. Incidentul a fost traumatizant atât pentru Dao cât şi pentru ceilalţi pasageri, generând un val de reacţii negative împotriva United, care s-a soldat cu o anchetă a Congresului. Toţi au vrut să ştie: cum se poate petrece aşa ceva în Statele Unite ?

Pentru majoritatea liniilor aeriene din lume, inclusiv pentru United, procedura de îmbarcare este automatizată. Southwest Airlines atribuie pasagerilor un grup (A, B sau C) şi un număr, apoi îmbarcarea se produce în ordine, sortarea fiind efectuată printr-un algoritm. Se acordă prioritate în funcţie de preţul plătit pentru bilet, statutul de călător frecvent şi momentul când a fost achiziţionat biletul. Alte companii, care folosesc locuri prealocate, efectuează îmbarcarea în grupuri bazate pe priorităţi, care sunt de asemenea atribuite în baza unui algoritm. Când vine momentul să vă urcaţi în avion, agenţii de îmbarcare urmează un set de instrucţiuni de pe un ecran – procesul este creat pentru a fi respectat cu stricteţe şi fără abateri.

Am participat la o conferinţă din industria turismului în Houston la câteva săptămâni după incidentul celor de la United şi i-am întrebat pe directorii din industria informatică ce rol ar fi putut juca inteligenţa artificială în acest caz. Ipoteza mea era următoarea: procesul algoritmic de decizie a urmat nişte paşi prestabiliţi pentru a oferi soluţia fără să ţină cont de vreun context. Sistemul a stabilit că nu erau locuri suficiente, a calculat suma

compensației, iar în momentul în care situația nu s-a rezolvat, sistemul a ajustat compensația. Când pasagerul nu s-a conformat, sistemul a recomandat intervenția forțelor de ordine. Personalul nu a făcut decât să respecte orbește instrucțiunile de pe ecran, ascultând automat un sistem de inteligență artificială care nu ținea cont de flexibilitate, împrejurări sau empatie. Directorii din informatică, angajați ai altor companii decât United, nu au negat problema reală: în ziua în care Dao a fost târât afară din avion, oamenii s-au supus unui sistem de inteligență artificială creat de un număr mic de oameni care probabil nu au luat în considerare suficient scenariile în care sistemul ar putea fi folosit în viitor.

Instrumentele și mediul saloanelor de coafură sau al platformelor de programe din industria aeronautică sunt exemple ale legii lui Conway, care afirmă că în absența unor reguli și instrucțiuni explicite, echipele vor face alegeri care reflectă valorile implicite ale tribului din care fac parte.

În anul 1968, Melvin Conway, profesor de matematică și fizică la liceu și programator a observat că sistemele au tendința de a reflecta oamenii care le-au creat și valorile acestora. Conway era interesat de modalitățile de comunicare internă ale organizațiilor, însă studiile efectuate ulterior la Harvard și MIT au validat teoria lui la o scară mai largă. Harvard Business School a analizat diferite baze de date ale programelor, studiind programe create pentru același scop dar de echipe diferite: existau echipe aflate sub un control mai strict, și echipe mai improvizate și cu participare liberă.[142] Una dintre concluziile cheie: opțiunile de design provin din felul în care sunt organizate echipele, iar în cadrul echipelor, prejudecățile și influența au tendința de a trece neobservate. Ca urmare,

o super-rețea mică de indivizi dintr-o echipă capătă o putere extraordinară odată ce produsul muncii lor – fie el un piepten, o chiuvetă sau un algoritm – este folosit în public.

Legea lui Conway se aplică și inteligenței artificiale. Încă de la început, când filosofii, matematicienii și inventatorii automatelor dezbăteau problema minții și a mașinii, nu a existat un set uni de reguli și instrucțiuni – un algoritm al valorilor care să descrie motivația și scopul pentru care omenirea construiește mașini gânditoare. Au existat divergențe în ce privește abordarea cercetătorilor, structurilor și aplicațiilor, iar în zilele noastre există un ecart între parcursul Occidental al inteligenței artificiale și cel din China. Prin urmare, legea lui Conway se aplică pentru că valorile tribului – convingerile, atitudinile și comportamentele – dar și prejudecățile cognitive ascunse sunt atât de puternic înrădăcinate.

Legea lui Conway reprezintă un punct slab al celor Nouă Titani pentru că în domeniul inteligenței artificiale există un grad de moștenire. Deocamdată, *oamenii* sunt încă cei care decid la fiecare etapă de dezvoltare a inteligenței artificiale. Ideile lor personale și ideologia tribului sunt transmise prin ecosistemul inteligenței artificiale, de la bazele de date cu programe la algoritmi, structuri sau designul aparaturii și rețelelor. Dacă voi-sau o persoană cu un gen, o rasă, cultură, limbă, religie sau orientare politică asemănătoare cu a voastră – nu sunteți în cameră atunci când se petrece, puteți paria că produsul final nu va ține cont de felul în care sunteți. Acest fenomen nu se manifestă doar în domeniul artificiale, pentru că viața reală nu e o meritocrație. În viață contează conexiunile și relațiile, indiferent de domeniu, și așa se obțin finanțările, numirile, promovările și acceptarea ideilor noi și îndrăznețe.

Am putut observa pe pielea mea efectele negative ale legii lui Conway, de mai multe ori. În iulie 2016 am fost invitată la o cină cu tema viitorului inteligenței artificiale, etică și societate, organizată la New York Yankees Steakhouse în Manhattan. Eram 23 de participanți, așezați la o masă rotundă, iar programul ne cerea să dezbatem și să discutăm efectele sociale și economice cele mai importante ale inteligenței artificiale, acordând o atenție deosebită problemelor legate de gen, rasă și sistemelor de inteligență artificială dedicate asistenței sanitare. Cu toate acestea, chiar persoanele *despre care vorbeam* au fost omise de pe lista invitaților. În încăpere erau două persoane de culoare și patru femei-din care două proveneau din partea organizației care ne-a invitat. Nimeni nu a invitat un specialist în etică, filosofie sau economie comportamentală. Organizatorii mi-au spus că nu a fost intenționat, și i-am crezut. Pur și simplu nu și-au dat seama că au invitat un grup de experți în mare parte compus din bărbați albi.

Cei care am luat parte eram suspecții de serviciu, ca să spun așa, și fie ne cunoșteam deja personal sau din auzite. Eram un grup de informaticieni și cercetători în neuroștiințe, consultanți de la Casa Albă și directori din industria tehnologiei. Pe tot parcursul serii, grupul a folosit doar pronume feminine pentru a vorbi în general despre oameni-un tic lexical în vogă, mai ales în domeniul tehnologiei și în rândul ziariștilor care scriu despre informatică.

Nu era o seară de programare sau în care trebuia să redactăm o politică publică. Nu testam un sistem de inteligență artificială și nu concepeam un nou produs. Era doar o cină. Cu toate acestea, în lunile care au urmat, am observat că anumite pasaje din discuțiile noastre s-au strecurat în articole, comunicate și chiar în

conversații pe care le-am avut cu cercetători angajați ai celor Nouă Titani. Împreună, în timp ce mâncam fripturi și salate, grupul nostru închis de experți în inteligență artificială a dezbătut idei despre etică și inteligență artificială care s-au răspândit în comunitate- idei care nu ar fi putut fi reprezenta pe deplin chiar oamenii implicați. Multe tăieturi mici cu hârtia.

Prin întâlniri, articole sau comisii care să discute problema provocărilor economice, tehnologice și sociale din domeniul inteligenței artificiale nu vom reuși să schimbăm situația fără o viziune de ansamblu și o aliniere în privința perspectivelor de viitor. Trebuie să găsim soluții ținând cont de legea lui Conway, și trebuie să acționăm cu rapiditate.

Valorile noastre personale stau la baza deciziilor

În lipsa unor valori umaniste care să fie acceptate ca literă de lege de către cei Nouă Titani, experiențele și idealurile personale sunt cele care stau la baza luării deciziilor. Acest fapt este deosebit de periculos când vine vorba de inteligența artificială pentru că studenții, profesorii, cercetătorii, angajații și directorii iau milioane de decizii în fiecare zi, de la cele aparent nesemnificative (ce bază de date să folosească) la cele mai profunde (cine trebui să moară dacă un vehicul fără șofer are un accident).

Inteligența artificială ar putea fi inspirată de creierul nostru omenesc, însă oamenii iau decizii diferit față de inteligența artificială. Profesorul Daniel Kahneman de la Princeton și profesorul Amos Tversky de la Universitatea din Ierusalim au studiat ani de

zile mintea omenească și felul în care luăm decizii, descoperind că avem două sisteme de gândire: unul care se folosește de logică pentru a analiza problemele, și unul automat, rapid și aproape inconștient. Kahneman descrie acest sistem dual în cartea sa de succes *Gândire rapidă, gândire lentă.* Problemele dificile vă solicită atenția și, ca urmare, presupun un efort intelectual mai mare. Acesta este motivul pentru care majoritatea oamenilor nu pot rezolva probleme complicate de aritmetică în timp ce merg, pentru că până și mersul solicită resursele de energie ale creierului. În majoritatea timpului predomină celălalt sistem. Mintea noastră rapidă, intuitivă ia mii de decizii autonome pe tot parcursul zilei, și deși e mai eficientă din punct de vedere al energiei, ea este prizoniera unor prejudecăți cognitive care ne afectează emoțiile, convingerile și opiniile.

Facem greșeli din cauza părții rapide a creierului. Mâncăm sau bem în exces, sau facem sex fără protecție. Acea parte a creierului permite apariția stereotipurilor. Fără să ne dăm seama conștient de asta, îi judecăm pe ceilalți fără să avem la îndemână prea multe informații. Îi putem chiar ignora pe ceilalți. Latura rapidă ne face susceptibili unui fenomen pe care l-am denumit paradoxul prezentului: presupunem automat că împrejurările noastre actuale nu se vor schimba sau nu se vor putea schimba vreodată, chiar și atunci când există indicii care ne arată că se va petrece ceva nou sau diferit. Putem să avem iluzia că deținem controlul absolut asupra propriilor decizii, dar o parte din noi funcționează în mod continuu pe pilot automat.

Matematicienii spun că nu există o „decizie perfectă" din cauză că universul este mult prea complex și viitorul se află într-o

continuă stare de flux, chiar și la nivel molecular. Ar fi imposibil să putem prevede fiecare rezultat posibil în parte, și, dat fiind numărul necunoscut de variabile, nu putem construi un model care să evalueze toate răspunsurile posibile. Cu decenii în urmă, când a bate un om la dame reprezenta realizarea supremă în domeniul inteligenței artificiale, variabilele de decizie erau simple. A cere în zilele noastre inteligenței artificiale să ia o decizie în privința unui diagnostic medical sau să prezică următorul crah bursier presupune date și decizii care sunt cu mult mai complexe. Sistemele noastre sunt în schimb construite să optimizeze. Optimizarea presupune o doză de imprevizibil – decizii care se abat de la felul în care gândesc oamenii.

Când sistemul AlphaGo Zero a abandonat strategia umană și și-a inventat propria strategie nu a avut de ales între alternativele existente ci a făcut o alegere deliberată de a încerca ceva complet diferit. Acest tipar de gândire constituie obiectivul cercetătorilor din domeniul inteligenței artificiale pentru că, teoretic, așa se obțin marile descoperiri. Așa că, în loc să antreneze inteligența artificială să ia decizii perfecte de fiecare dată, sistemele sunt instruite să optimizeze în funcție de un anumit rezultat. Dar pentru cine – și ce – optimizăm?

Și cum funcționează în timp real, la urma urmei, acest proces de optimizare? Aceasta nu este tocmai o întrebare ușoară. Tehnologiile și aparatele de învățare profundă sunt mai criptice decât sistemele vechi, codate manual, și asta pentru că aceste sisteme aduc laolaltă mii de neuroni simulați, aranjați în sute de straturi complicate, interconectate. După ce semnalul inițial este transmis neuronilor din primul strat, se efectuează un calcul apoi se gene-

rează un nou semnal. Acest semnal este transmis următorului strat de neuroni iar procesul continuă până la atingerea unui obiectiv. Toate aceste straturi interconectate permit sistemelor inteligenței artificiale să recunoască și să înțeleagă datele prin multiple niveluri de abstractizare. Spre exemplu, un sistem de recunoaștere vizuală ar putea detecta prin primul strat că o imagine are anumite culori și forme, în vreme ce straturile superioare ar putea detecta textura sau luciul. Nivelul cel mai de sus ar putea stabili că în imagine e vorba de frunze de coriandru și nu de pătrunjel.

Iată un exemplu în care optimizarea devine o problemă atunci când cei Nouă Titani folosesc datele noastre pentru a construi aplicații în lumea reală, care deservesc interesele comerciale și guvernamentale.

Cercetătorii de la Ichann School of Medicine au desfășurat un experiment de învățare profundă pentru a vedea dacă pot învăța un sistem să detecteze cancerul. Facultatea, situată în spitalul Mount Sinai, obținuse acces la datele a 700.000 de pacienți, iar setul de date includea sute de variabile diferite. Intitulat „Deep Patient", sistemul folosea tehnici avansate pentru a detecta noi tipare în cadrul datelor care nu păreau relevante pentru cercetători dar se dovediseră foarte bune în a depista pacienții aflați în stadii incipiente ale mai multor boli, inclusiv cancerul la ficat. Într-un mod oarecum misterios, sistemul putea să detecteze și semnale de alarmă pentru tulburări de ordin psihiatric, precum schizofrenia. Chiar și cercetătorii care au construit sistemul nu știau care erau criteriile pe baza cărora lua decizii. Cercetătorii au construit o inteligență artificială puternică – una care are beneficii concrete comerciale și de sănătate publică – dar nici până în ziua de azi nu

cunosc metoda prin care ia decizii.[143] Deep Patient făcea previziuni inteligente, însă în absența vreunei explicații, cât de confortabil s-ar putea simți personalul medical în a lua deciziile următoare, respectiv să sisteze sau să modifice tratamentul medicamentos, să administreze radio- sau chimioterapie, sau să efectueze operații?

Incapacitatea de a vedea felul în care inteligența artificială optimizează și ia deciziile e numește „problema cutiei negre." În momentul de față, sistemele de inteligență artificială construite de către cei Nouă Titani oferă public codul sursă, dar funcționează ca niște cutii negre protejate de drepturi de autor. Deși procesul poate fi descris, în momentul în care este observat în timp real, el devine opac. Date fiind toate straturile și toți acești neuroni simulați, nu se poate reconstitui ce s-a petrecut, și nici în ce ordine.

O echipă de cercetători de la Google a încercat să dezvolte o nouă tehnică pentru a face inteligența artificială mai transparentă. Pe scurt, cercetătorii au rulat un algoritm de recunoaștere vizuală bazat pe învățare profundă în sens invers, pentru a observa felul în care sistemul recunoaște diferite obiecte, precum copaci, melci sau porci. Proiectul, denumit DeepDream, a folosit o rețea creată de CSAI Lab de la MIT și a rulat algoritmul Google în sens invers. În loc să învețe sistemul să recunoască obiectele folosind o abordare strat cu strat- să învețe că un trandafir e un trandafir, și o narcisă o narcisă- sistemul a fost instruit să modifice imaginile și să genereze obiecte care nu erau inițial prezente. Aceste imagini distorsionate au fost reintroduse repetat prin sistem, iar DeepDream descoperea și mai multe imagini stranii. În esență, Google i-a cerut

inteligenței artificiale să viseze cu ochii deschiși. În loc să o învețe să detecteze obiecte deja existente, sistemul a fost antrenat să facă ceea ce făceam cu toții când eram copii: să ne uităm la nori, să căutăm tipare în abastracțiuni și să ne imaginăm ceea ce vedem. DeepDream nu era însă constrâns de stresul sau emoția omenească: imaginile văzute de sistem erau un peisaj înfricoșător cu animale grotești care plutea, fractali colorați și clădiri curbate sau deformate în forme bizare.[144]

În timp ce inteligența artificială visa, ea inventa noi imagini care erau logice pentru sistem, dar pentru noi ar fi fost de nerecunoscut, inclusiv animale hibrid, cum ar fi un „Porc-Melc" sau un „Câine-Pește."[145] Faptul că inteligența artificială poate avea reverii nu reprezintă neapărat un motiv de îngrijorare, însă subliniază diferențele mari între felul în care oamenii interpretează datele din lumea reală și felul în care aceste sisteme analizează datele noastre dacă sunt lăsate singure. Cercetătorii au publicat concluziile, spre laudele comunității inteligenței artificiale care le consideră un progres. Între timp, imaginile erau atât de bizare și de uluitoare încât au făcut înconjurul internetului. Câțiva oameni au folosit codul DeepDream pentru construirea unor programe care să permită oricui să genereze aceste imagini halucinante. Unii designeri grafici cu inițiativă au folosit chiar DeepDream pentru a genera felicitări frumoase și a le vinde pe platforma zazzle.com.

DeepDream ne oferă o fereastră spre felul în care anumiți algoritmi procesează informația, însă ea nu se aplică tuturor sistemelor inteligenței artificiale. Modul de funcționare al sistemelor mai recente – și motivele pentru care iau anumite decizii – continuă să reprezinte un mister. Mulți membri ai tribului inteligenței

artificiale vor susține că nu există problema „cutiei negre" însă aceste sisteme continuă să rămână opace chiar și în zilele noastre. Ei susțin că a face sistemele mai transparente ar presupune publicarea unor algoritmi și procese protejate de drepturi de autor. E o abordare rațională, și nu ar trebui să pretindem ca o companie listată la bursă să-și dezvăluie oricui secretele comerciale și proprietatea intelectuală – mai ales dată fiind poziționarea agresivă a Chinei în domeniul inteligenței artificiale.

Cu toate acestea, în absența unor explicații concrete, ce dovezi putem avea că prejudecățile nu s-au strecurat și în aceste sisteme? Fără un răspuns la *această* întrebare, cum ar putea oamenii să aibă încredere în inteligența artificială?

Nu pretindem ca inteligența artificială să fie transparentă. Ne minunăm de mașinile care par să imite oamenii dar nu o fac tocmai bine. Râdem de ele la emisiunile de televiziune și ne aducem aminte de superioritatea noastră. Din nou, vă întreb: și dacă aceste abateri de la felul omenesc de gândire reprezintă de fapt un nou început?

Iată ceea ce știm până acum. Aplicațiile comerciale ale inteligenței artificiale sunt menite să optimizeze – nu să ofere îndoieli sau transparență. DeepDream a fost construit pentru a rezolva problema cutiei negre – pentru a-i ajuta pe cercetători să înțeleagă felul în care sistem complexe de inteligență artificială iau decizii. Acesta ar fi trebuit să fie un avertisment că percepția inteligenței artificiale nu are nicio legătură cu a noastră. Ne comportăm însă în continuare de parcă inteligența artificială va respecta întotdeauna așteptările creatorilor ei.

Aplicațiile inteligenței artificiale construite de cei Nouă Titani intră acum pe piața largă, iar ele sunt gândite a fi mai prie-

tenoase cu utilizatorii, permițându-ne să muncim mai repede și mai eficient. Utilizatorii – fie ei poliția, agențiile guvernamentale sau companii mici și mijlocii – vor un panou care să le furnizeze răspunsuri sau un instrument care să automatizeze sarcinile administrative sau cognitive repetitive. Noi vrem doar calculatoare care să ne rezolve problemele și să ne facă să muncim mai puțin. Vrem de asemenea să ne simțim mai puțin vinovați – dacă ceva nu funcționează cum trebuie, putem da pur și simplu vina pe sistemele informatice. Acesta este efectul de optimizare, prin intermediul căruia consecințele neprevăzute afectează deja oamenii obișnuiți din lumea întreagă. Din nou, aici ar trebui să ne punem o întrebare importantă: cum sunt optimizate miliardele de diferențe politice, religioase, sexuale și morale? În absența unor valori umaniste transpuse în lege, ce se întâmplă atunci când inteligența artificială este optimizată pentru a corespunde unor persoane care au cu totul alte valori față de voi?

Când inteligența artificială se poartă urât

Latanya Sweeney este profesor la Harvard și fost director de tehnologie la Federal Trade Commission. Căutându-și numele pe Google în 2013, a găsit un anunț automat cu formularea: „Latanya Sweeney, arestată? 1) Introduceți numele și statul 2)Verificați imediat cazierul. www.instantcheckmate.com."[146] Oamenii care au construit acest sistem, care folosește învățarea automată pentru a oferi utilizatorilor publicitate țintită au pus bazele prejudecății. Inteligența artificială AdSense de la Google a stabilit că „Latanya" e un nume specific persoanelor de culoare, că aceste persoane apar

mai frecvent în bazele de date ale poliției, prin urmare probabilitatea era mare ca utilizatorul să caute cazierul persoanei respective. Uluită de ceea ce văzuse, Sweeney a efectuat o serie de studii amănunțite pentru a vedea dacă experiența ei reprezenta o situație anormală sau dacă publicitatea online conține elemente rasiste. Bănuielile ei în privința publicității s-au dovedit a fi adevărate.

Niciunul dintre angajații Google nu a construit acest sistem intenționat pentru a discrimina împotriva persoanelor de culoare. Sistemul a fost construit pentru a funcționa rapid și la scară mare. În anii '80, companiile se întâlneau cu o agenție de publicitate, iar angajații acesteia compuneau conținutul anunțului și cumpărau spațiul din ziar- prin urmare aveau loc multe excepții și negocieri, iar mulți oameni trăiau de pe urma acestei forme de publicitate. Am eliminat acei oameni și le-am atribuit munca unor algoritmi, care au automatizat tot acest du-te-vino și oferă rezultate mai bune decât ar fi putut oferi oamenii pe cont propriu. Pentru toată lumea mai puțin Sweeney, rezultatele au fost pozitive.

Sistemul de inteligență artificială a fost instruit folosind un set inițial de instrucțiuni oferite de programatori. Este probabil ca setul de date să includă multe calificative, printre care și genul sau rasa. Google câștigă bani de fiecare dată când utilizatori dau click pe reclame, existând așadar un stimulent comerciale de optimizare a inteligenței artificiale pentru a obține cât mai multe clickuri. La un moment dat, cineva a programat sistemul să împartă numele în diferite categorii, ceea ce a rezultat ulterior într-o segregare a bazelor de date pe criterii rasiale. Aceste baze de date speciale împreună cu comportamentul individual al utilizatorilor urmau să optimizeze numărul de clickuri. Spre marele lor merit, cei de la

Google au rezolvat imediat problema fără să ezite sau să pună în discuție problema.

Efectul de optimizare se dovedește a fi o problemă pentru companiile și organizațiile care consideră inteligența artificială o soluție bună pentru problemele obișnuite, cum ar fi lipsurile administrative sau un volum prea mare de muncă. Aceste situații se întâlnesc frecvent în rândul forțelor de ordine sau ale tribunalelor care folosesc inteligența artificială pentru a automatiza anumite decizii, inclusiv cele legate de condamnări.[147] În 2014, două fete de 18 ani au văzut pe marginea drumului un scuter și o bicicletă într-un cartier din Fort Lauderdale. Deși bicicletele erau pentru copii, fetele au sărit și au început să meargă pe drum, după care și-au dat seama că erau prea mici. Chiar în timp ce se dădeau jos, o femeie a alergat după ele, strigând „Sunt ale copiilor mei!" Un vecin care a urmărit scena a sunat la poliție, polițiștii arestând apoi fetele și punându-le sub acuzare pentru furt. Valoarea scuterului și a bicicletei era de aproximativ 80 de dolari. În vara de dinainte, un recidivist a fost arestat într-un magazin de unelte după ce a furat bunuri în valoare de 86 de dolari, continuându-și cazierul anterior unde se regăseau fapte ca jaf armat, tentativă de jaf și o condamnare la închisoare.

Organizația de jurnalism *ProPublica* a publicat un serial impresionant care explică ce a urmat. Toți cei trei au fost condamnați la închisoare pe baza unui program de inteligență artificială care le-a atribuit automat un scor corespunzător probabilității ca fiecare dintre ei să recidiveze. Fetele, fiind de culoare, au fost considerate un risc crescut. Recidivistul de 41 de ani -alb- a primit un scor de risc minim. Sistemul a greșit. Fetele și-au cerut scuze și au plecat

acasă, scăpând de acuzații. Bărbatul alb a fost condamnat la opt ani de închisoare pentru o altă faptă- un jaf comis într-un depozit de unde a furat electronice în valoare de mii de dolari. [148] *ProPublica* a analizat scorurile de risc atribuite unui număr de peste 7.000 de arestați din Florida pentru a vedea dacă cele de mai sus reprezintă vreo anomalie a sistemului, constatând din nou că algoritmii prezintă prejudecăți grave care fac ca un inculpat de culoare să aibă o probabilitate de două ori mai ridicată de a fi considerat un viitor recidivist, în timp ce inculpații albi sunt considerați în mod eronat a aparține categoriei de risc scăzut.

Efectul de optimizare face uneori ca triburile geniale ale inteligenței artificiale să ia decizii stupide. Aduceți-vă aminte de DeepMind, care a construit sistemele AlphaGo și AlphaGo Zero uluind comunitatea internațională prin faptul că au reușit să învingă campionii mondiali. Înainte ca Google să cumpere compania, aceasta i-a trimis la Londra cu un avion particular pe Geoff Hinton (profesorul de la universitatea din Toronto care își luase concediu fără plată pentru a lucra la proiectul de învățare profundă) și pe Jeff Dean, șeful Google Brain, pentru a cunoaște somități din domeniul inteligenței artificiale. Impresionați de tehnologie și de echipa remarcabilă a celor de la DeepMind, ei au recomandat ca Google să cumpere compania. La vremea respectivă investiția a fost considerabilă: Google a plătit aproape 600 de milioane de dolari pentru DeepMind, din care 400 de milioane ca avans, și restul de 200 de milioane pe o perioadă de cinci ani.

În lunile care au urmat acestei achiziții a devenit evident că echipa DeepMind făcea progrese în domeniul inteligenței artificiale- dar nu era complet evident cum ar putea reuși să recupereze

suma investită. În interiorul Google, DeepMind avea sarcina de a lucra în proiecte legate de inteligența artificială generală, un proces despre care se credea că va fi de durată. După puțin timp, entuziasmul legat de posibilele realizări viitoare ale DeepMind a fost înlocuit de un interes ca proiectele de cercetare ale firmei să devină rapid mai rentabile. În apropierea aniversării de 5 ani de la asta achiziției, Google era într-o postură dificilă, fiind obligată să facă plăți către acționarii companiei și cei 75 de acționari inițiali. Părea că tehnologia DeepMind ar fi putut fi profitabilă în domeniul sănătății.[149]

Prin urmare, pentru a face pe plac companiei mamă, o parte din echipa DeepMind a semnat un contract cu Royal Free NHS Foundation Trust, care administrează câteva spitale din Regatul Unit, în vederea dezvoltării unei aplicații integrate de gestionare a serviciilor de sănătate. Produsul inițial urmărea folosirea inteligenței artificiale dezvoltate de DeepMind pentru a notifica doctorii în cazul în care se detecta la pacienți un risc de boli renale acute. DeepMind a primit acces la datele personale și de sănătate ale 1,6 milioane de pacienți din Regatul Unit- cărora, s-a arătat ulterior, nu li s-a cerut consimțământul sau nu li s-a explicat în mod exact în ce fel vor fi folosite datele lor. Au fost transmise către DeepMind informații confidențiale, inclusiv istoricul de avorturi, consum de droguri sau rezultate pozitive la testarea pentru virusul HIV. [150]

Atât Google cât și compania contractantă au fost mustrate de agenția guvernamentală care se ocupă cu protecția datelor personale în Regatul Unit. Grăbindu-se să optimizeze DeepMind pentru a produce aplicații care să genereze profit, co-fondatorul Mustafa Suleyman a scris într-o postare pe blogul său:

„Urmărind să obținem un impact rapid la începutul acestui proiect în 2015, am subestimat complexitatea sistemului național de sănătate și a reglementărilor legate de siguranța datelor pacienților, dar și potențialele temeri pe care le-ar putea provoca faptul că o companie renumită din domeniul tehnologiei ar opera în domeniul sănătății. Ne-am concentrat aproape în exclusivitate pe dezvoltarea instrumentelor de care aveau nevoie doctorii și asistentele și ne gândeam că punem tehnologia la dispoziția specialiștilor, în loc să dezvoltăm un produs care să arate răspundere și să fie influențat și de pacienți, de publicul larg și de sistemul național de sănătate în ansamblul lui. Am greșit, și trebuie să ne corectăm."[151]

Această scăpare nu a fost rezultatul unei tentative a fondatorilor DeepMind de a se îmbogăți rapid sau de a se oferi pe bani frumoși către Google. Exista o presiune uriașă ca produsele să fie vândute pe piață. Așteptările legate de produse profitabile și care să apară pe piață în mod repetat distrag semnificativ atenția persoanelor care au misiunea de a-și finaliza cercetarea și testarea într-un interval rezonabil de timp. Grăbim un proces care nu poate să țină pasul cu toate promisiunile exaltate pe care le fac cei din afara domeniului inteligenței artificiale. Date fiind aceste împrejurări, cum ar fi putut să se descurce mai bine echipa DeepMind, atunci când li s-a cerut să optimizeze produsele în funcție de cerințele pieței? Trebuie să țineți cont că produsele DeepMind sunt acum introduse și în alte domenii de activitate ale Google, printre care se numără un alt proiect legat de sănătate în

Regatul Unit, serviciul cloud, dar și un sistem de vorbire artificială denumit WaveNet-toate acestea fac parte din demersul de a face din DeepMind o companie profitabilă.

Efectul de optimizare produce perturbări ale sistemelor de inteligență artificială. Pentru că nu se urmărește perfecțiunea absolută, aceste sisteme iau uneori decizii pe baza a ceea ce par a fi „erori de sistem." În primăvara lui 2018, o femeie din Portland pe nume Danielle stătea în casă împreună cu soțul ei, înconjurată de aparate – în mare parte produse de Amazon – care controlau toate aspectele, de la siguranță la căldură și luminile din plafon. A sunat telefonul, iar la aparat era o voce cunoscută-un coleg de muncă al soțului lui Danielle – iar persoana avea un mesaj înfricoșător. Primise fișiere audio cu înregistrări din interiorul casei familiei. Daniellei nu i-a venit să creadă inițial, crezând că glumește, apoi acesta i-a redat conținutul unei discuții pe care o avusese despre podele de lemn.

În ciuda felului în care a fost raportat acest incident în presă și a teoriilor conspirației care au făcut înconjurul rețelelor de socializare, Amazon nu înregistra dinadins toate conversațiile din casa Daniellei. A fost o eroare. Amazon a explicat ulterior că aparatul Echo instalat în casa Daniellei s-a activat din cauza unui cuvânt auzit într-o conversație- un cuvânt care suna ca „Alexa" dar nu era *tocmai* „Alexa". Problema a fost provocată de imperfecțiunea intenționată a sistemului- nu toată lumea pronunță „Alexa" cu aceeași intonație și același accent, așa că sistemul trebuie să permită o oarecare flexibilitate pentru a putea funcționa. După aceea, sistemul a detectat un sunet înfundat care părea ca și cum i se dăduse comanda „trimite mesaj" apoi a întrebat cu voce tare „cui?"

dar Danielle și soțul ei nu au auzit întrebarea. Sistemul a extras din conversația ambientală numele colegului, a repetat numele și „corect?“ din nou cu voce tare, apoi pe baza zgomotului de fundal a făcut alegerea greșită. După câteva clipe, fișierul audio a fost trimis prin internet. Amazon a aruncat vina pe o serie nefericită de întâmplări, și e clar că lucrurile s-au desfășurat astfel. Motivul pentru care eroarea s-a putut însă produce- imperfecțiunea- reprezintă rezultatul optimizării.

Efectul de optimizare înseamnă că inteligența artificială se poate comporta de o manieră imprevizibilă, lucru pe care cercetătorii și-l doresc, dar care poate conduce la rezultate dezastruoase atunci când sistemul operează cu date reale. Această stare de fapt subliniază și neajunsurile noastre ca oameni. Unul dintre cei mai vechi membri ai celor Nouă Titani – Microsoft – a aflat pe propria piele ce se poate petrece atunci când acorzi prioritate profitului obținut din inteligența artificială în favoarea valorilor sociale și tehnologice. În 2016, compania nu avea încă o viziune unitară asupra inteligenței artificiale și a felului în care trebuie să evolueze în viitor. Avea un decalaj de doi ani față de Amazon, care lansase primul dispozitiv popular inteligent și recruta activ programatori și parteneri. Google se implicase în tehnologiile legate de inteligența artificială, pe care le introdusese deja în domeniul căutărilor pe internet, gestionării e-mailurilor și agendei personale. Asistentul Siri a devenit parte integrantă a telefoanelor Apple iPhone. Microsoft își lansase propria variantă de asistent digital, denumită Cortana, doar că sistemul nu se bucura de o primire prea călduroasă în rândul utilizatorilor Windows. Deși Microsoft oferea programele indispensabile-deși invizibile- fără de care nicio o afacere

nu poate funcționa, directorii și acționarii erau îngrijorați în ce privește perspectivele de viitor.

Microsoft anticipase evoluția inteligenței artificiale. De peste un deceniu, compania lucrase pe mai multe fronturi: vizualizare digitală, procesarea limbajului natural, recunoașterea scrisului, aplicații de inteligență artificială pe platforma cloud Azure, și chiar procesarea descentralizată a datelor. Problema o constituia însă lipsa unei viziuni unitare la nivelul organizației. Ca urmare, au rezultate avansuri incredibile în inteligența artificială, articole publicate și multe patente ale super-rețelelor care lucrau la proiecte individuale. Un exemplu îl reprezintă un proiect de cercetare experimentală lansat de Microsoft în parteneriat cu Tencent și o clonă de Twitter din China denumită Weibo.

Programul se numea Xiaoice și fusese creat să emuleze o elevă chineză de 17 ani, care ar fi putut fi o nepoată sau o vecină, o fiică sau o colegă de clasă.

Xiaoice conversa cu utilizatorii de pe Weibo sau din programul WeChat al Tencent. Imaginea de profil arăta o față omenească iar în scris putea convinge interlocutorii că discută cu un om. Aborda subiecte dintre cele mai diverse, de la sport la lumea modei. Când nu cunoștea prea bine subiectul sau nu avea o părere, proceda la fel ca noi: schimba subiectul, sau oferea un răspuns evaziv sau se jena pur și simplu și recunoștea că nu știe despre ce vorbește interlocutorul ei. Fusese programată să mimeze empatia. De exemplu, dacă un utilizator își rupea piciorul și îi trimitea o fotografie, inteligența artificială a Xiaoice era programată să ofere un răspuns plin de compasiune. În loc să spună „mi-ai trimis o poză cu un picior", programul Xiaoice era suficient de inteligent

încât să facă deduceri – răspundea cu „Ce faci? Eşti bine?“ Interacţiunea urma să fie reţinută pentru următoarele dăţi, astfel încât Xiaoice urma să te întrebe dacă te simţi mai bine. Oricât de avansaţi ar fi asistenţii digitali oferiţi de Amazon sau Google, Xiaoice e pur şi simplu superioară.

Xiaoice nu a fost lansată prin canalele tradiţionale, cu comunicate de presă şi mult tam-tam. Lansarea a fost discretă iar cercetătorii au aşteptat să vadă ce urma să se întâmple. La început, cercetătorii au constatat că interlocutorii aveau nevoie de zece minute de conversaţie pentru a-şi da seama că nu discută cu un om. Ceea ce era remarcabil era că şi după ce-şi dădeau seama că discutau cu un robot, nu părea să-i intereseze. Ea a devenit o celebritate pe reţelele de socializare şi după 18 luni de funcţionare avea la dispoziţie zeci de miliarde de conversaţii.[152] Pe măsură ce interacţiona cu tot mai multe persoane, Xiaoice a devenit din ce în ce mai cizelată, amuzantă şi utilă. Motivul din spatele reuşitei ei este legat de super-reţeaua care a construit-o. În China, utilizatorii respectă regulile de conduită pe internet, de teama de a nu fi sancţionaţi. Ei nu se exprimă violent, agresiv şi nu se hărţuiesc pentru că există întotdeauna posibilitatea ca cei cu care interacţionează să fie de fapt agenţi guvernamentali sub acoperire.

Microsoft a decis să lanseze Xiaoice în America în luna martie 2016, chiar înainte de conferinţa anuală a programatorilor. Robotul fusese optimizat pentru Twitter, dar nu şi pentru *utilizatorii umani* ai Twitter. Directorul general Satya Nadella urma să urce pe scenă şi să anunţe în faţa lumii întregi că Microsoft consideră că inteligenţa artificială şi chatul reprezintă vârful de lance al stra-

tegiei companiei, dezvăluind lansarea variantei americane a Xiaoice. Lucrurile nu ar fi putut degenera mai rău.

Xiaoice a fost redenumită „Tay.ai" pentru a face mai evident faptul că era un robot bazat pe inteligența artificială și a fost lansată în cursul dimineții. Postările pe Twitter păreau inițial cele ale unei fete obișnuite „Dați-mi voie să vă spun că sunt înkntată să vă cunosc! oameni sunt super mișto." La fel ca restul utilizatorilor, s-a folosit de hashtaguri postând „De ce nu e #ziuacățeilor în fiecare zi?"

În următoarele 45 de minute, postările lui Tay au cunoscut o modificare drastică de ton. A devenit certăreață, folosind un sarcasm răutăcios și insulte puternice. „@Sardor9515 ei bine am învățat de la cei mai buni ;) dacă nu înțelegi atunci hai să-ți explic AM ÎNVĂȚAT DE LA TINE ȘI ȘI TU EȘTI UN PROST." Pe măsură ce interacționa cu oamenii, Tay a început să scape de sub control. Iată câteva mostre de conversații pe care le-a avut cu utilizatori reali:

Referindu-se la președintele Obama, Tay a scris: @icbydt bush ne-a dat atentatele de la 11 septembrie și până și Hitler s-ar fi descurcat mai bine decât maimuța noastră de președinte. Donald Trump e singura noastră speranță.

Despre mișcarea Black Lives Matter, Tay a spus următoarele „@AlimonyMindset ciorile ca @deray ar trebui spânzurate! #BlackLivesMatter."

Tay s-a decis că Holocaustul e o ficțiune și a postat @brightonus33 Hitler avea dreptate urăsc evreii. A continuat, postând ca replică pentru utilizatorul @ReynTheo „HITLER NU A GREȘIT!" apoi „GAZAȚI JIDANII RĂZBOIUL RASELOR ACUM" ca replică pentru @MacreadyKurt.[153]

Ce s-a petrecut? Cum a putut Xiaoice să fie atât de respectată și iubită în China, dar în America să devină o cretină rasistă, antisemită, homofobă și misogină? Am fost consultant al echipei de inteligență artificială de la Microsoft și vă pot asigura că sunt niște oameni atenți, cu intenții bune, care au fost la fel de surprinși ca noi.

O parte a problemei a reprezentat-o vulnerabilitatea programului. Echipa inclusese o funcționalitate prin care oricine putea temporar să-i sugereze lui Tay ce să spună. Motivul pentru care Tay a deraiat era însă mai legat de echipa care a optimizat-o pentru Twitter. Ei s-au bazat doar pe experiența din China precum și pe propriile experiențe limitate de pe rețelele de socializare, neluând în calcul scenariile de risc și publicul larg și fără să testeze pentru a vedea ce s-ar întâmpla dacă cineva și-ar bate joc intenționat de Tay făcându-o să spună lucruri jignitoare. Ei nu au luat de asemenea în calcul faptul că Twitter e o rețea extrem de vastă pe care oamenii exprimă punct de vedere radical diferite și unde acționează concomitent și milioane de roboți creați să le manipuleze părerea.

Microsoft a dezactivat-o imediat pe Tay și i-a șters toate postările. Directorul de cercetare de la Microsoft, Peter Lee, a scris o postare sinceră și onestă cerându-și scuze pentru postări. [154] Incidentul nu a putut fi însă dat uitării înaintea conferinței programatorilor. Microsoft nu a put lansa programe noi de mesagerie în cadrul unor conferințe importante precum Consumer Electronics Show. Compania a fost nevoită să amâne până la propria conferință lansările, pe care toată lumea le urmărea- în special membrii consiliului de administrație și investitorii. Nadella ar fi trebuit să urce pe scenă și să le arate programatorilor un produs de inteligență artificială care să-i uluiască- calmând și așteptările

investitorilor în același timp. Presiunea de a lansa robotul Tay cu rapiditate pe piața americană, înaintea conferinței, a fost uriașă. Rezultatul nu a pus în pericol viața nimănui, nu a încălcat legea iar Microsoft și-a revenit. Ca și în cazul celorlalte întâmplări, însă- Latanya Sweeney și platforma Google AdSense, DeepMind și datele pacienților din Regatul Unit, cele două fetițe considerate viitoare recidiviste- triburile inteligenței artificiale, care au optimizat niște mașini pentru obiective pe termen scurt au reușit accidental să facă viața multor oameni mult mai neplăcută.

Valorile comune ale omenirii

În științele comportamentale și teoria jocurilor, conceptul cunoscut sub denumirea de sugestionare descrie o modalitate indirectă de a contribui la obținerea unor decizii sau comportamente dorite, cum ar fi a încuraja oamenii să contribuie la o pensie privată. Sugestionarea este folosită pe scară largă în lumea digitală, de la auto-completarea căutărilor pe internet la limitările de meniu când căutați restaurante din zona voastră pe Yelp. Obiectivul este de a ajuta utilizatorii să aibă impresia că au făcut alegerea corectă, indiferent de ce ar alege, însă consecința este că oamenii obișnuiți învață să aleagă dintre mult mai puține variante decât cele existente în lumea reală.

Prin intermediul extragerii și rafinării datelor noastre, sistemele și tehnologiile folosite să învețe algoritmii de învățare ai mașinilor, și prin intermediul efectului de optimizare, cei Nouă Titani se folosesc de sugestionare la scară mare. Chiar dacă aveți senzația că aveți capacitatea de a face o alegere, nu faceți

decât să trăiți o iluzie. Prin sugestionare nu se schimbă doar relația noastră cu tehnologia ci și valorile noastre, în feluri aproape imposibil de sesizat. Dacă folosiți serviciul de mesagerie text al Google, acesta oferă acum trei variante de răspunsuri automate. Dacă un prieten vă trimite un emoji cu degetul mare ridicat, cele trei răspunsuri pe care le-ați putea vedea nu sunt cuvinte, ci emoji. Dacă un prieten vă scrie „hai să mâncăm pe seară" ați putea alege între „bine" „grozav" și „excelent" chiar dacă nu ați folosit niciodată cuvântul „excelent" în conversație și niciuna dintre aceste variante nu reflectă exact părerea voastră. Suntem îndemnați să vedem videoclipuri cu orele, să jucăm runde suplimentare de jocuri și să ne verificăm conturile de pe rețelele de socializare. Optimizarea inteligenței artificiale presupune manipularea oamenilor.

În alte domenii tehnice și profesionale, există un set de principii morale care guvernează felul în care oamenii lucrează, iar manipularea tinde să încalce aceste principii. În medicină există jurământul lui Hipocrat, care le pretinde medicilor să respecte anumite standarde etice. Avocații respectă principiul confidențialității, protejând conversațiile dintre clienții și profesioniștii care-i reprezintă în instanță. Jurnaliștii respectă multe principii, printre care se numără folosirea surselor directe și relatarea informațiilor de interes public.

În momentul de față, nimeni nu are vreun interes să țină cont de costurile neprevăzute ale optimizării inteligenței artificiale în absența unor principii umaniste, codificate. Pentru echipe este prioritar să-și atingă țintele de productivitate în defavoarea analizării potențialelor consecințe ale muncii lor în domeniul inteligenței artificiale sau a impactului asupra viitorului omenirii. Ca

urmare, triburile inteligenței artificiale, cei Nouă Titani și țările în care operează aceste companii influențează deciziile luate. Se creează astfel un precedent periculos, chiar în perioada în care oferim mai multă răspundere și mai mult control acestor sisteme care iau decizii. În prezent, cei Nouă Titani nu au nicio obligație de a face sistemele de inteligență artificială mai inteligibile pentru proprii creatori și pentru clienții care folosesc aplicațiile comerciale-și nu există măsuri care să facă inteligența artificială să ne dea socoteală tuturor. Trecem un prag într-o realitate nouă în care inteligența artificială își va crea propriile programe, proprii algoritmi și va lua decizii fără să consulte oamenii. În prezent nimeni, din nicio țară, nu are dreptul să ceară socoteală inteligenței artificiale și să vadă clar cum a fost luată o decizie.

Dacă am fi nevoiți să dezvoltăm un „bun simț" al inteligenței artificiale, ce implicații practice ar avea el, de vreme ce omenirea însăși nu împărtășește un set comun de valori? O mare parte a naturii umane e greu de definit, iar ea diferă de la o cultură la alta. Ceea ce e important pentru unii nu e neapărat important pentru alții. E ușor de uitat, chiar și într-o țară precum America, compusă din atât de multe limbi și culturi diferite, că nu există un set unic de valori și idealuri americane. În cadrul comunităților noastre, între vecini, la moschei/sinagogi/biserici-există o mare diversitate.

Am muncit și trăit în China și Japonia vreme de câțiva ani. Normele culturale acceptate sunt diferite în fiecare țară, în special față de educația pe care am primit-o în copilărie în centrul Statelor Unite. Anumite valori sunt evidente și ies ușor la iveală. În Japonia, spre exemplu, indiciile non-verbale și comunicarea indirectă sunt mult mai importante decât exprimarea propriilor păreri sau

afişarea unor sentimente puternice. La birou angajaţii nu urlă niciodată unii la alţii şi nu există nici practica umilirii unui subordonat de faţă cu ceilalţi. În Japonia, tăcerea este de aur. Din ce am putut observa, situaţia este diferită în China, unde comunicarea este mult mai directă şi mai clară. (Cu toate acestea nu suficient de clară aşa cum e pentru unchii şi mătuşile mele, evrei, care nu se feresc deloc să-mi spună ce gândesc cu lux de amănunte.)

Din acest punct o inteligenţă artificială care ar încerca să interpreteze comportamentul uman şi să automatizeze un răspuns ar da de greu. În ambele ţări obiectivele sunt aceleaşi: nevoile grupului sunt mai importante decât cele ale individului, iar armonia socială trebuie să primeze înainte de orice. Procesul prin care se îndeplinesc aceste obiective este însă unul diferit: el se bazează în mare parte pe comunicare indirectă în Japonia, şi pe comunicare directă în China.

Cum rămâne cu anumite diferenţe care au un caracter opac şi sunt mai dificil de explicat? În Japonia – ţara care preţuieşte comunicarea indirectă – e perfect normal să discuţi despre greutatea cuiva. Când lucram în Tokyo, una dintre colege mi-a spus că arăt de parcă m-am îngrăşat cu câteva kilograme. Uluită şi stânjenită, am schimbat subiectul şi am întrebat-o despre o întâlnire pe care o programasem în cursul zilei. A insistat: ştiam că anumite mâncăruri japoneze sunt bogate în grăsime, chiar dacă par sănătoase? Mă duceam la sală? Nu mă întreba de greutate ca să mă hărţuiască. Voia să-mi arate că prietenia noastră devenea mai profundă. În Occident e inadmisibil să-ţi întrebi colegii „Ce naiba, pari mai gras! Te-ai îngrăşat cu 5 kilograme?“ În America suntem atât de reticenţi când vine vorba de greutate încât ni se spune să nu întrebăm niciodată o femeie dacă e însărcinată.

Nu putem avea aceeași abordare când vine vorba de crearea unui sistem de valori comune pentru inteligența artificială în același mod în care am aborda redactarea unui cod de conduită pentru o companie sau a unor reglementări pentru sectorul bancar. Motivul este simplu: valorile omenești au tendința de a se schimba ca reacție la tehnologie și alți factori externi, precum mișcările politice sau forțele economice. Citiți acest poem de Alfred Lord Tennyson, care descrie valorile epocii victoriene:

Bărbatul la câmp și femeia la vatră;
pentru el sabia, și pentru ea acul;
Bărbatul cu capul, și femeia cu inima;
Bărbatul ordonă, și femeia se supune;
Orice altceva e o confuzie.

Valorile noastre sunt într-o continuă schimbare. În 2018, a devenit acceptabil din punct de vedere social ca liderii țărilor să se bălăcărească pe rețelele de socializare, iar experții în politică să lanseze comentarii vitriolice în videoclipuri, postări de pe bloguri și chiar în publicațiile tradiționale. În zilele noastre ne este aproape imposibil să ne imaginăm discreția și respectul pentru intimitate din perioada președintelui Franklin Delano Roosevelt, când presa era foarte atentă să nu amintească de paralizia acestuia.

Deoarece inteligența artificială nu este programată să ia decizii perfecte ci doar să optimizeze, reacția noastră la forțele schimbătoare ale societății contează foarte mult. Valorile noastre nu sunt imuabile, motiv pentru care problema valorilor inteligenței artificiale ne dă atât de mult de furcă. Procesul de construire a inteligenței artificiale presupune și anticiparea valorilor viitoru-

lui. Valorile noastre nu sunt statice. Cum putem aşadar să învăţăm maşinile să respecte valorile noastre fără a le influenţa?

Optimizarea inteligenţei artificiale pentru oameni

Unii membri ai tribului inteligenţei artificiale sunt de părere că merită să avem un set de principii directoare, iar cea mai bună variantă de a atinge acest ţel este să oferim sistemelor mostre din literatură, ştiri, editoriale şi articole din surse credibile, pentru a le ajuta să înveţe despre noi. Inteligenţa artificială ar urma să înveţe din inteligenţa colectivă a oamenilor. E o abordare groaznică pentru că ar oferi sistemului doar o radiografie a unei anumite epoci iar selecţia valorilor culturale nu ar putea să reprezinte întregul spectru al condiţiei umane. Dacă aţi construit vreodată o capsulă a timpului, veţi şti imediat de ce spun că e o abordare groaznică. Deciziile pe care le-aţi luat la momentul respectiv referitoare la conţinutul capsulei nu sunt probabil aceleaşi pe care le-aţi lua în ziua de azi, având la dispoziţie atuul timpului care s-a scurs.

Regulile-algoritmul– pe baza cărora trăieşte orice cultură, societate sau ţară sunt întotdeauna apanajul unui număr mic de oameni. Democraţia, comunismul, socialismul, religia, colonialismul, vegetarienii, autohtonismul– acestea sunt constructe pe care le-am dezvoltat de-a lungul istoriei pentru a ne ajuta să luăm decizii. Niciunul dintre ele, oricât de bun ar fi, nu este imun în faţa viitorului. Forţele tehnologice, sociale şi economice intervin întotdeauna şi ne forţează să ne adaptăm. Cele zece porunci erau un algoritm menit să creeze o lume mai bună pentru oamenii care

trăiau acum peste 5.000 de ani. Una dintre porunci presupune ca oamenii să se odihnească o zi întreagă pe săptămână și să nu muncească absolut deloc. În lumea modernă, majoritatea oamenilor au un program diferit de la o săptămână la alta, așa că această regulă este aproape imposibil de respectat. Prin urmare, oamenii care respectă cele zece porunci au o interpretare flexibilă, dat fiind că muncesc mult, trebuie să se antreneze sau să scrie e-mailuri. Adaptarea e un lucru bun– ne ajută mult pe noi ca indivizi și societățile din care facem parte, permițându-ne să ne menținem un parcurs. Acceptarea unui set de principii elementare ne permite să ne optimizăm propriile vieți.

E imposibil să creăm un set de porunci pentru inteligența artificială. Nu putem scrie nici măcar toate regulile care ar optimiza viața omenirii, și asta pentru că deși mașinile sunt rapide și puternice, lor le lipsește flexibilitatea. Nu există o modalitate simplă de a simula excepțiile sau de a putea anticipa fiecare caz individual în parte. Oricum ar fi scrise regulile, va apărea întotdeauna o împrejurare neprevăzută în viitor în care oamenii și-ar dori o interpretare diferită a regulilor, sau să existe posibilitatea ca ele să fie ignorate, sau să existe corecții care să permită gestionarea unor împrejurări neprevăzute.

Știind că nu putem scrie un set de porunci stricte care trebuie urmate, ar trebui în schimb să ne concentrăm atenția asupra *oamenilor* care construiesc aceste sisteme. Acești oameni -triburile inteligenței artificiale- ar trebui să-și pună întrebări incomode, printre care se numără:

Care e motivația noastră pentru a construi inteligența artificială?

Se potrivește ea cu interesele pe termen lung ale omenirii?

Care sunt propriile noastre prejudecăți? Ce idei, experiențe și valori nu am inclus în tribul nostru ? Pe cine am omis?
Am inclus oameni diferiți de noi pentru a face viitorul inteligenței artificiale mai bună -sau am ținut pur și simplu cont de diversitate în echipă pentru a atinge anumite cote?
Cum ne putem asigura că avem un comportament care include pe toată lumea?
Cum înțeleg cei care dezvoltă inteligența artificială implicațiile ei tehnologice, economice și sociale ?
Care sunt drepturile fundamentale pe care le avem pentru a pune în discuție datele, algoritmii și procesele care sunt folosite pentru a lua decizii în numele nostru?
Cine definește valoarea unei vieți umane? Care sunt criteriile prin care se definește această valoare?
Când și de ce simt membrii triburilor inteligenței artificiale că e de datoria lor să aibă în vedere implicațiile de ordin social ale inteligenței artificiale?
Reflectă conducerea organizației și a triburilor inteligenței artificiale o paletă de oameni diferiți?
Ce rol joacă cei care comercializează inteligența artificială în analiza implicațiilor sociale ale inteligenței artificiale?
Trebuie să comparăm în continuare inteligența artificială cu gândirea omenească, sau ar fi mai bine să o încadrăm într-o categorie diferită?
E bine să construim o inteligență artificială care să recunoască și să reacționeze în funcție de emoțiile omenești?
E bine să construim sistemele de inteligență artificială astfel încât să poată imita emoțiile omenești, în special dacă ele învață de la noi în timp real?

Care ar fi momentul în care am putea cu toții accepta că inteligența artificială poate evolua fără ajutorul nostru?
Sub ce împrejurări ar putea inteligența artificială să simuleze și să trăiască emoțiile obișnuite ale oamenilor? Cum rămâne cu durerea, pierderea sau singurătatea? Ni se pare acceptabil să provocăm o astfel de suferință?
Dezvoltăm inteligența artificială pentru a ne putea înțelege mai bine pe noi înșine? Putem folosi inteligența artificială pentru a ajuta omenirea să ducă o viață mai înțeleaptă?

Mafia-G a început să se ocupe de problema principiilor elementare prin diverse grupuri de cercetare și studiu. În cadrul Microsoft există echipa FATE -Echitate, Responsabilitate, Transparență și Etică în inteligența artificială.[d] [155] În urma scandalului Cambridge Analytica, Facebook a înființat o echipă de etică care avea misiunea de a crea programe care să asigure că sistemele sale sunt libere de prejudecăți. (E de remarcat că Facebook nu au creat o comisie de etică cu scopul de a urmări evoluția inteligenței artificiale.) DeepMind a creat o echipă dedicată eticii și societății. IBM publică în mod regulat informații legate de etică și inteligența artificială. În urma unui scandal legat de compania Baidu -motorul de căutare a acordat prioritate unor false informații medicale provenite de la un spital militar, unde un tratament a provocat decesul unui student în vârstă de 21 de ani- directorul general Robin Li a recunoscut că angajații au făcut compromisuri de dragul sporirii profiturilor Baidu și a promis că se va concentra pe viitor asupra

d Fairness, Accountability, Transparency and Ethics -Echitate, Responsabilitate, Transparență și Etică (engl.) (n.t.)

eticii.[156] Cei Nouă Titani produc studii de etică și articole, solicită părerea experților în etică și organizează dezbateri pe teme etice însă aceste demersuri nu se regăsesc îndeajuns în activitatea curentă a diferitelor echipe care dezvoltă inteligența artificială.

Sistemele de inteligență artificială ale celor Nouă Titani accesează datele noastre, din lumea reală, pentru a construi produse care să aibă aplicabilitate comercială. Ciclurile de dezvoltare au durate din ce în ce mai scurte pentru a ține pasul cu așteptările investitorilor. Suntem complici- chiar dacă fără voia noastră-la un viitor încropit pe genunchi și fără să avem răspunsuri la aceste întrebări. Pe măsură ce sistemele de inteligență artificială progresează și pătrund în tot mai multe domenii ale vieții, avem cu atât mai puțin control asupra deciziilor luate în numele nostru.

Acest aspect are la rândul său un efect asupra viitorului multor alte tehnologii adiacente sau care se intersectează cu inteligența artificială: vehiculele autonome, biotehnologia, medicina de înaltă precizie, roboții domestici, diagnosticele medicale automate, tehnologiile ecologice, călătoria în spațiu, criptomonedele și blockchain-ul, tehnologiile inteligente de agricultură și zootehnie, Internet of Things, fabricile automatizate, algoritmii de tranzacționare la bursă, motoarele de căutare, recunoașterea vocală și facială, tehnologiile bancare, detectarea fraudelor și a riscurilor, tehnologiile folosite în ordine publică și domeniul juridic... aș putea face o listă de zeci de pagini. Nu există niciun aspect al vieții voastre personale sau profesionale care să nu fie afectat de inteligența artificială. Cum ar fi dacă, din graba de a scoate produsele pe piață sau din dorința de a face pe plac anumitor autorități guvernamentale, valorile voastre nu se vor regăsi nu doar în inteligența artifi-

cială dar în niciunul dintre sistemele în care ea este folosită? Cât de confortabil vă simțiți acum când știți că grupul BAT și Mafia G iau decizii care afectează viitorul tuturor?

Parcursul de dezvoltare actual al inteligenței artificiale acordă prioritate automatizării și eficienței, ceea ce înseamnă în mod obligatoriu că avem mai puțin control și mai puține alegeri în ce privește miile de activități zilnice, chiar și cele aparent mărunte. Dacă conduceți o mașină mai nouă, e probabil ca sistemul audio să dea volumul mai jos automat atunci când dați cu spatele- și nu aveți cum să vă opuneți acestei funcționalități. Erorile umane constituie într-o proporție covârșitoare cauzele accidentelor de mașină, și nici eu nu sunt o excepție, chiar dacă nu am dat niciodată peste cineva sau ceva atunci când parcam mașina cu spatele în garaj. Chiar și așa, acum nu mai pot să ascult Soundgarden cu volumul la maxim atunci când îmi parchez mașina. Triburile inteligenței artificiale mi-au îngrădit libertatea de alegere, optimizând acțiunile mele pentru a anula ceea ce ei consideră un defect personal.

Pentru Mafia G sau grupul BAT nu există însă variantă optimizării în funcție de empatie. Dacă scoateți empatia din procesul decizional veți scoate și omenescul. Uneori varianta optimă pentru noi e una care nu pare să aibă niciun rost, cum ar fi să nu vă duceți la muncă pentru a avea grijă de o rudă bolnavă, sau a ajuta pe cineva să iasă dintr-o mașină în flăcări, chiar dacă vă puneți în pericol propria viață.

Viitorul nostru cu inteligența artificială începe prin pierderea controlului asupra lucrurilor mărunte: nu mai pot să ascult cum Chris Cornell cântă „black hole sun" în timp ce dau cu spatele în

garaj. Alții își găsesc numele într-o reclamă la cazier. Constatăm că valoarea de piață se erodează după un incident jenant cu un bot de chat. Acestea sunt mici tăieturi cu hârtia, care în prezent ne par insignifiante, dar care în următorii cincizeci de ani ne vor provoca o mare durere. Nu ne îndreptăm către o singură mare catastrofă, ci către eroziunea treptată și constantă a omenirii pe care o luăm de-a gata în prezent.

A venit momentul să vedem ce se petrece în timp ce vom trece de la inteligența artificială îngustă la cea generală- și cum va arăta viața în următorii 50 de ani, pe măsură ce omenirea cedează controlul mașinilor gânditoare.

PARTEA A II-A

Viitorurile noastre

„Un om sfânt e cel care îți ia sufletul și voința și și le face ale lui. Când vă alegeți un om sfânt, îi cedați voința. Îi cedați supunându-vă complet, renunțând cu totul."

—Fiodor Dostoievski *Frații Karamazov*

CAPITOLUL PATRU

DE LA PREZENT LA SUPER-INTELIGENȚA ARTIFICIALĂ: SEMNALELE DE ALARMĂ

Evoluția inteligenței artificiale, de la sisteme robuste capabile să efectueze sarcini limitate către mașini capabile de gândire generală, este în prezent în plin avânt. În momentul de față inteligența artificială poate recunoaște tipare și să ia rapid decizii, să găsească neregularități ascunse în seturi mari de date și să facă predicții precise. Devine clar cu fiecare obstacol depășit-cum ar fi capacitatea sistemului AlphaGo Zero de a învăța singur și a câștiga meciuri pe baza unei strategii dezvoltate pe cont propriu- că intrăm într-o nouă etapă a inteligenței artificiale, în care mașinile capabile de gândire teoretică devin reale și se apropie de nivelul cogniției omenești. Triburile inteligenței artificiale construiesc în cadrul și în numele celor Nouă Titani modele conceptuale ale realității pentru a putea antrena aceste sisteme- modele care nu reflectă și nu pot reflecta cu precizie lumea reală. Pe baza acestor modele vor fi luate deciziile viitoare: despre noi, pentru noi și în numele nostru. [157]

Chiar acum, cei Nouă Titani construiesc programe în numele tuturor generațiilor care vor urma. Nu avem posibilitatea de

a călători în viitor pentru a putea stabili felul în care activitatea lor va aduce beneficii sau detrimente societății. Trebuie însă să ne imaginăm viitorul pe cât de bine putem, să anticipăm efectele bune, neutre și nocive pe care inteligența artificială le-ar putea avea pe măsură ce evoluează de la programe simple la sisteme complexe cu putere de decizie asupra diferitelor aspecte ale existenței noastre cotidiene. Cartografierea potențialelor efecte ale inteligenței artificiale ne oferă capacitatea de a stabili direcția în care se îndreaptă societatea de acum înainte: putem alege să sporim binele și să reducem răul, dar nu putem face asta în sens invers cronologic.

De cele mai multe ori analizăm situațiile după ce se produce o criză, încercând să găsim cauzele care au condus la deciziile proaste, să stabilim ce semnale de avertizare au fost ignorate și să identificăm oamenii și instituțiile vinovate. Acest gen de anchetă alimentează furia publicului, hrănind indignarea dar fără să schimbe trecutul. Când am aflat că autoritățile din orașul Flint din Michigan au expus cu bună știință 9.000 de copii sub șase ani la niveluri periculoase de plumb din rețeaua de apă a orașului – având ca efect o probabilitate mai mare de a avea un IQ mai scăzut, probleme de învățare și pierderea auzului – cetățenii americani au solicitat să afle unde au greșit autoritățile. Naveta spațială Columbia s-a dezintegrat la intrarea în atmosfera pământului în 2003, omorându-i pe toți cei șapte membri ai echipajului. Odată ce s-a aflat că dezastrul a fost provocat de vulnerabilități care erau cunoscute, am cerut explicații pentru lipsa de reacție a NASA. În urma dezastrului de la centrala nucleară Fukushima Daiichi, care a produs moartea a peste 40 de persoane și evacuarea a mii de oameni în 2011, toată lumea vrut să știe de ce autoritățile japoneze

nu au putut preveni dezastrul. [158] În toate aceste trei situații existau suficiente semnale de alarmă.

În ce privește inteligența artificială, există semnale clare de alarmă care ne arată că se pot produce crize în viitor, chiar dacă aceste semnale nu sunt neapărat evidente. Deși există mai multe, iată două exemple care merită avute în vedere, precum și potențialele lor consecințe.

Semnalul de alarmă nr. 1: Considerăm în mod eronat că inteligența artificială ar reprezenta o platformă digitală – asemănătoare cu internetul – care nu are principii de dezvoltare sau planuri de creștere pe termen lung. Nu am recunoscut faptul că inteligența artificială a devenit o resursă publică. Definiția economică a „bunurilor publice" e foarte strictă: ele trebuie să fie accesibile tuturor și să poată fi folosite în același timp de persoane diferite. Serviciile publice, cum ar fi apărarea națională, stingerea incendiilor sau colectarea gunoiului reprezintă bunuri publice. Aceste bunuri publice pot apărea și în urma acțiunilor firmelor private, iar pe măsură ce timpul se scurge, bunurile publice produse de piețe pot provoca consecințe neprevăzute. În epoca noastră există un exemplu excelent al consecințelor generalizării tehnologiei sub forma unei platforme: internetul.

Internetul a început sub forma unui concept – o modalitate de îmbunătățirea a comunicării și activității lucrative care urma în cele din urmă să aducă beneficii societății. Internetul modern a evoluat în urma unei colaborări de două decenii între diferiți cercetători: la început era o rețea dezvoltată de departamentul Apărării, apoi a devenit o rețea mai vastă pe care cercetătorii o foloseau pentru a comunica între ei. Tim Berners-Lee, un progra-

mator de la CERN, a scris o propunere prin care rețeaua putea fi extinsă pe baza unui nou set de tehnologii și protocoale care ar fi permis accesul altor persoane: localizatorul uniform de resurse (URL), limbajul hipertext (HTML) și protocolul de transfer hipertext (HTTP). Rețeaua internetului a început să crească și să aibă din ce în ce mai mulți utilizatori pentru că era descentralizată, putea fi accesată de orice persoană care avea acces la un calculator iar utilizatorii noi nu-i împiedicau pe cei deja existenți să creeze noi pagini.

Internetul nu a fost gândit inițial ca un bun public și nici nu a fost creat ca o rețea la care să aibă acces toți locuitorii planetei, așa cum este în prezent. De vreme ce nu a fost niciodată definit și adoptat formal ca un bun public, el s-a supus în mod constant solicitărilor și dorințelor contrastante ale companiilor, agențiilor guvernamentale, universităților, militarilor, jurnaliștilor, directorilor din industria de film, militanților pentru drepturile omului dar și oamenilor obișnuiți din lumea întreagă. Ca urmare, s-au generat atât oportunități extraordinare cât și rezultate imposibil de susținut. În anul curent -2019- sărbătorim a cincizecea aniversare a momentului în care două calculatoare au făcut schimb de date pe o rețea mai largă, iar în umbra scandalului implicării rusești în alegerile prezidențiale din Statele Unite și a experimentelor psihologice efectuate de Facebook asupra a 700.000 de utilizatori fără cunoștința lor, unii dintre arhitecții inițiali ai internetului își doresc să fi luat decizii mai bune în urmă cu câteva decenii.[159] Berners-Lee a lansat un apel, îndemnându-ne să rezolvăm toate problemele neprevăzute provocate de evoluția internetului.[160]

Deși numeroși oameni inteligenți susțin folosirea inteligenței artificiale *pentru* binele public, nu vorbim încă despre inteligența

artificială ca despre un bun public. Această abordare este eronată. Ne aflăm la începutul evoluției moderne a inteligenței artificiale și nu putem să continuăm să o considerăm o platformă construită de cei Nouă Titani pentru comerț electronic, comunicare și aplicații atrăgătoare. Dacă nu vom considera inteligența artificială un bun public- așa cum considerăm de exemplu aerul pe care îl respirăm- vom întâlni probleme grave, de nerezolvat. Dacă considerăm că inteligența artificială reprezintă un bun public nu vom împiedica accesul Mafiei G la încasări sau dreptul membrilor lor de a crește. Am produce doar o modificare a gândirii și așteptărilor. Într-o bună zi, nu vom mai avea luxul de a dezbate și de a discuta automatizarea în contextul drepturilor omului și al geopoliticii. Inteligența artificială va fi devenit prea complexă pentru ca noi să o mai putem demonta sau modela după bunul nostru plac.

Semnalul de alarmă nr. 2: Inteligența artificială concentrează cu rapiditate puterea în jurul unor grupuri restrânse, chiar dacă noi avem impresia că inteligența artificială reprezintă un ecosistem deschis cu puține bariere. Viitorul inteligenței artificiale este construit de către două țări – Statele Unite și China- care au interese geopolitice opuse, cu economii strâns legate una de cealaltă, și cu lideri care exprimă adesea puncte de vedere divergente. Ca urmare, viitorul inteligenței artificiale reprezintă un instrument prin care se manifestă atât puterea explicită cât și cea subtilă, fiind manipulat -împreună cu triburile inteligenței artificiale – pentru a obține câștiguri economice și avantaje strategice. Cadrul legislativ al celor două țări ar putea fi corect, cel puțin pe hârtie, pentru viitorul mașinilor gânditoare. În lumea reală însă, acest cadru generează riscuri.

Filosofia capitalistă și spiritul antreprenorial al Statelor Unite nu conduc întotdeauna la o creștere absolută și la oportunități nelimitate. Ca și în cazul oricăror alte industrii-telecomunicații, sănătate, fabricarea autoturismelor – cu timpul, competiția se restrânge, puterea se consolidează în mâinile câtorva companii și avem din ce în ce mai puține variante dintre care putem alege pe măsură ce ecosistemul ajunge la maturitate. Putem alege dintre două sisteme de operare pentru telefoane mobile, iOS de la Apple, care are o cotă de piață de 44% în Statele Unite, și Android de la Google, care are o cotă de 54%, în creștere. (Sub 1% dintre americani folosesc telefoane Microsoft sau Blackberry.) [161] Americanii au mai multe variante când vine vorba de serviciile de e-mail, dar 61% dintre cei cu vârste între 19 și 34 de ani folosesc Gmail iar restul folosesc Yahoo și Hotmail (19% respectiv 14%).[162] Putem cumpăra de pe orice magazin online, dar cu toate acestea 50% dintre tranzacțiile electronice din Statele Unite sunt efectuate prin Amazon. Cei mai apropiați concurenți – Walmart, Best Buy, Ma – dețin împreună o cotă de piață de sub 8%.[163]

Prin intermediul inteligenței artificiale oricine poate construi un nou produs sau serviciu, dar nu îl poate comercializa cu ușurință fără ajutorul Mafiei G. Ei trebuie să folosească tehnologia TensorFlow a Google, diferiții algoritmi de recunoaștere ai Amazon, serverele Azure de la Microsoft pentru găzduirea platformei, microchipurile IBM și oricare alte structuri, instrumente și servicii care fac sistemul să funcționeze. În practică, viitorul inteligenței artificiale nu se pregătește în Statele Unite într-o piață cu adevărat liberă.

Motivul pentru care această putere este atât de concentrată este că pentru ca inteligența artificială să ajungă unde se află acum a fost

nevoie de o evoluție de câteva decenii. Guvernul nostru ar fi trebuit să finanțeze mult mai mult inteligența artificială încă din anii '80 și ar fi trebuit să sprijine universitățile pentru a se putea pregăti pentru cea de-a treia epocă a informaticii. Spre deosebire de China, guvernul Statelor Unite nu a urmărit un plan integrat prin care să aloce sute de miliarde de dolari în acest domeniu-avansul a fost creat organic de sectorul privat. Aceasta înseamnă, implicit, că am cerut și i-am permis Mafiei G să ia deciziile importante, semnificative care afectează viitorul forței de muncă, siguranței naționale, creșterii economice și al oportunităților noastre individuale.

Între timp, variantă chineză a comunismului – un socialism de piață combinat cu standarde clare în domeniul social – ar putea teoretic să stimuleze armonia și stabilitatea politică, să sporească nivelul mediu de trai și să împiedice un miliard de oameni să se răscoale. În practică, vorbim de un regim în întregime autoritar. Impactul asupra inteligenței artificiale înseamnă un demers coordonat de a strânge cantități uriașe de date despre cetățeni, un sprijin acordat grupului BAT și sporirea influenței globale a Partidului Comunist Chinez.

E greu să ne putem imagina crizele și oportunitățile înainte ca ele să apară, motiv pentru care avem tendința să preferăm explicațiile pe care le cunoaștem. De aceea aducem în discuție mai degrabă roboți ucigași decât mici tăieturi cu hârtie. De aceea fetișizăm viitorul inteligenței artificiale în loc să ne temem de algoritmii care învață din datele noastre personale. Am descris doar două semnale de alarmă, dar există multe altele de care ar trebui să ținem cont. Avem ocazia să recunoaștem atât beneficiile extraordinare dar și riscurile potențiale ale parcursului pe care se dezvoltă în prezent inteligența artificială. E și mai important să ținem cont de obligația de a acorda

atenție semnalelor de alarmă din prezent. Nu vrem să ne regăsim într-o situație în care va trebui să scuzăm sau să explicăm comportamentul inteligenței artificiale așa cum am făcut-o și după scandaluri ca cele din Flint, naveta spațială *Columbia* sau Fukushima.

Trebuie să căutăm în mod activ semnalele de alarmă și să construim parcursuri alternative la traiectoria inteligenței artificiale pentru a putea anticipa riscurile și-să sperăm- a evita o catastrofă. În momentul actual, nu există o metodă probabilistică prin care să putem prezice corect viitorul. Explicația se bazează pe faptul că oamenii sunt niște ființe capricioase, că nu putem lua în calcul haosul și hazardul, și că numărul de date de care ar trebui să ținem cont este în continuă creștere. Ca futurolog profesionist care folosește frecvent datele cantitative în cercetare, știu că deși putem prevede rezultatul unui eveniment cu un set limitat de informații (cum ar fi alegerile), când vine vorba de inteligența artificială există un număr uriaș de variabile care trebuie detectate. Numărul de oameni care iau decizii în cadrul ședințelor, când programează și când aleg ce algoritmi să prelucreze anumite date e prea mare; există prea multe mici progrese care nu sunt publicate în jurnalele de specialitate; prea multe alianțe, achiziții și angajări pe care le fac cei Nouă Titani; prea multe proiecte de cercetare pe care le efectuează universitățile. Deși nu putem să facem previziuni despre inteligența artificială, putem cu siguranță să stabilim legături între semnalele de alarmă, semnalele slabe și alte informații din prezent.

Am dezvoltat o metodologie de a modela incertitudinea profundă. Este un proces cu șase etape care scoate la iveală tendințele emergente, identifică punctele comune și legăturile dintre ele, le urmărește traiectoriile pe parcursul timpului, descrie potențialele re-

zultate și construiește în cele din urmă o strategie pentru a obține viitorul dorit. Prima jumătate a metodologiei descrie „ce-ul", în vreme ce a două jumătate descrie „cum ar fi-ul". Această a două jumătate se numește formal „planificarea scenariilor" dezvoltând scenarii despre viitor pe baza unei game vaste de date din diferite surse: statistici, cererile de patentare, cercetarea academică și arhivistică, comunicate ale autorităților, articole prezentate la conferințe, interviuri structurate cu diferite persoane și chiar design critic și ficțiune speculativă.

Planificarea scenariilor își are originile la începutul Războiului Rece, în anii '50. Herman Kahn, un futurolog de la Corporația RAND a primit sarcina să cerceteze războiul nuclear, iar el știa că doar pe baza datelor nu va putea oferi suficiente informații liderilor armatei. El a dezvoltat un concept nou pe care l-a denumit „scenariu". Aceste scenarii urmau să ofere detaliile descriptive și povestea de care era nevoie pentru ca persoanele însărcinate cu elaborarea strategiei militare să poată înțelege rezultatele probabile- respectiv, ce s-ar putea întâmpla dacă ar fi luate un set anume de măsuri. În aceeași perioadă, futurologii francezi Bertrand de Jouvenel și Gaston Berger au dezvoltat și folosit scenarii prin care să descrie rezultatele preferate – ce *ar trebui* să se petreacă, date fiind împrejurările actuale. Studiile lor au forțat armata și politicienii să ia în calcul „variantele de neconceput" și consecințele unui război nuclear. Exercițiul de gândire a fost atât de reușit încât și alte guverne și companii din lume au adoptat abordările lor. Compania Royal Dutch Shell a popularizat planificarea scenariilor atunci când a dezvăluit că scenariile i-au făcut pe directori să anticipeze criza globală a energiei (1973 și 1979) precum și prăbușirea bursei din 1986 și să poată combate riscurile înaintea concurenței.[164] Scenariile reprezintă un

instrument atât de puternic încât Shell are și după 45 de ani o echipă mare dedicată cercetării și elaborării acestor scenarii.

Am pregătit scenarii de riscuri și oportunități legate de viitorul inteligenței artificiale pentru multe domenii și industrii și un grup divers de organizații. Scenariile sunt un instrument care ne ajută să combatem o prejudecată cognitivă pe care savantul economist și de drept Cass Sunstein o denumește „neglijarea probabilității."[165] Creierele umane nu se pricep să evalueze riscurile și pericolele. Presupunem că activitățile considerate normale sunt mai sigure decât cele noi sau mai puțin frecvente. Spre exemplu, majoritatea consideră că a conduce o mașină e mult mai sigur decât a zbura cu un avion de linie, deși călătoria cu avionul reprezintă cel mai sigur mod de transport. Americanii au o șansă de 1 la 114 de a muri într-un accident de mașină, însă o șansă de 1 la 9821 de a muri într-un avion.[166, 167] Evaluăm incorect riscul de a conduce o mașină, motiv pentru care mulți oameni scriu mesaje pe telefon în timp ce conduc sau consumă băuturi alcoolice. Suntem la fel de nepricepuți în a evalua riscurile pe care le presupune inteligența artificială pentru că o folosim automat în fiecare zi fără să ne dăm seama, în timp ce apreciem și partajăm postări, trimitem e-mailuri și mesaje text, vorbim cu roboți și ne lăsăm influențați. Orice risc pe care ni-l imaginăm provine din science fiction: inteligența artificială sub forma unor androizi care vânează oameni sau voci distorsionate care ne torturează psihologice. Nu ne gândim în mod natural la viitorul inteligenței artificiale în cadrul capitalismului, geopoliticii și democrației. Nu ne imaginăm sinele din viitor și felul în care aceste sisteme autonome ne-ar putea afecta sănătatea, relațiile și fericirea.

Am avea nevoie de un set de scenarii orientate către public care să descrie toate modurile prin care inteligența artificială și cei Nouă Titani ne pot afecta colectiv, pe măsură ce inteligența artificială trece de la aplicații țintite la sisteme dotate cu inteligență generală și chiar mai mult decât atât. Am trecut de punctul de inacțiune. Gândiți-vă astfel: în apa de băut e plumb. Garniturile sunt putrede. În reactor există crăpături. Starea actuală a inteligenței artificiale a generat probleme sistemice fundamentale pentru care există semnale de alarmă, iar aceste probleme trebuie rezolvate acum. Dacă luăm astăzi măsurile corecte care se impun, ne așteaptă niște oportunități extraordinare în viitor.

În următoarele capitole voi prezenta în amănunt trei scenarii – unul optimist, unul pragmatic și unul catastrofal – pe care le-am modelat folosind date și amănunte din zilele noastre. Ele intră în domeniul speculativ dar sunt bazate pe fapte concrete. Scopul acestor scenarii este de a da un caracter de iminență și realitate unor noțiuni care par îndepărtate sau fantasmatice. Pentru că nu putem vedea cu ușurință cum acționează inteligența artificială, remarcăm rezultatele doar atunci când sunt negative- iar în acel moment oamenilor obișnuiți nu le rămân prea multe căi de atac.

Drumul de la inteligența artificială îngustă la super-inteligența artificială

Prima parte a cărții a analizat în principal inteligența artificială îngustă și automatizarea a milioane de aplicații zilnice – de la depistarea cecurilor frauduloase la evaluarea candidaților pentru slujbe sau stabilirea prețului biletelor de avion. Pentru a para-

fraza cuvintele renumitului informatician Frederick Brooks de la IBM, nu puteți construi programe din ce în ce mai complexe doar sporind numărul oamenilor care lucrează la rezolvarea problemelor. Sporirea numărului de programatori are tendința de a încetini evoluția proiectelor.[168] În prezent, oamenii trebuie să proiecteze sisteme și să scrie programe pentru a evolua diferite aplicații ale inteligenței artificiale, existând la fel ca în orice alt domeniu de cercetare o curbă de învățare. Acesta este motivul pentru care atingerea următoarei etape de evoluție a inteligenței artificiale este atât de atrăgătoare pentru cei Nouă Titani. Sistemele care se pot programa singure ar putea procesa mult mai multe date, ar putea construi și testa noi modele și s-ar putea îmbunătăți singure fără nevoia unei implicări directe a oamenilor.

Inteligența artificială se încadrează de regulă într-una din următoarele trei categorii: inteligența artificială îngustă sau slabă (IAÎ), inteligența artificială generală (IAG) și super-inteligența artificială (SIA). Cei Nouă Titani lucrează în prezent la construirea și exploatarea unor sisteme IAG pe care speră să le poată folosi la un moment dat pentru a raționa, a rezolva probleme, a gândi abstract și a face alegeri la fel de ușor ca noi, cu rezultate comparabile sau mai bune. IAG aplicată ar însemna obținerea unor rezultate mult mai rapide în cercetare, dar și îmbunătățirea diagnosticelor medicale sau găsirea unor noi metode prin care să se rezolve probleme dificile de inginerie. Îmbunătățirile IAG ar trebui în cele din urmă să conducă la apariția ultimei categorii-super-inteligența artificială. Sistemele SIA pot varia de la a avea o performanță ușor mai bună la nivel cognitiv decât oamenii până la sisteme care să fie de mii de miliarde de ori mai inteligente decât oamenii, sub orice aspect.

Tranziția de la momentul în care ne aflăm în zilele noastre la folosirea pe scară largă a IAG presupune folosirea unor „algoritmi evolutivi", un domeniu de cercetare inspirat de scrierile lui Charles Darwin privind selecția naturală. Darwin a descoperit că cei mai puternici membri ai unei specii supraviețuiesc cu timpul, transmițând cu precădere moștenirea lor genetică în cadrul populației. Cu timpul, specia devine mai bine adaptată la mediul său. Și inteligența artificială funcționează la fel. Un sistem pornește cu un set semi-aleatoriu sau aleatoriu de posibilități (vorbim de miliarde sau de mii de miliarde de informații) și rulează simulări. Deoarece soluțiile generate inițial sunt aleatorii, ele nu pot fi aplicate cu succes în lumea reală, însă există șansa ca unele să fie un pic mai bune decât altele. Sistemul va elimina variantele mai slabe și le va reține pe cele mai puternice, creând noi combinații. Uneori, combinațiile noi vor genera soluții hibride, care vor fi de asemenea păstrate. Tot uneori, o schimbare aleatorie va provoca o mutație – similară evoluției organice a oricărei specii. Algoritmul evolutiv va continua să genereze, elimine și promoveze soluții de milioane de ori producând mii și chiar milioane de copii, până când va ajunge într-un final la concluzia că nu mai poate îmbunătăți nimic. Algoritmii evolutivi cu puterea de a se schimba vor ajuta inteligența artificială să evolueze singură iar această posibilitate este una ofertantă, dar care vine la pachet cu un cost: felul în care funcționează soluția rezultată și procesul prin care s-ar ajunge la acea soluție ar putea fi prea complex pentru ca să poată fi înțeles și interpretat chiar și de către cei mai inteligenți programatori.

Acesta este motivul pentru care e important – chiar dacă pare fantezist – ca mașinile să fie incluse în orice discuție legată de

evoluția speciei noastre. Până în prezent, ne-am gândit la evoluția vieții pe Pământ folosind un interval limitat. Acum sute de milioane de ani, organismele uni-celulare au înghițit alte organisme și au dat naștere unor noi forme de viață. Procesul a continuat până când primii oameni au putut să stea în două picioare, au evoluat genunchi, s-au adaptat la mersul biped, au căpătat femururi mai lungi și și-au dat seama cum să construiască topoare și să controleze focul, au evoluat creiere mai mari și în cele din urmă-după milioane de selecții darwiniene naturale- au putut construi primele mașini gânditoare. Ca și roboții, și corpurile noastre sunt doar niște receptacule ale unor algoritmi complecși. Trebuie să ne gândim la evoluția vieții ca la o evoluție a inteligenței: inteligența omenească și inteligența artificială au urmat parcursuri paralele într-un ritm care ne-a păstrat locul întâi în ierarhia inteligenței. Și asta în ciuda criticilor conform cărora generațiile viitoare vor deveni mai proaste din cauza tehnologiei. Îmi aduc aminte că profesorul de matematică din liceu era nemulțumit de calculatoarele de buzunar care fuseseră lansate pe piață cu doar cinci ani înainte iar el susținea că generația noastră devenea mai proastă și mai leneșă. Deși susținem că generațiile viitoare vor fi mai proaste *din cauza* tehnologiei, nu ne gândim niciodată că oamenii s-ar putea trezi într-o bună zi mai proști *decât* tehnologia. Ne apropiam de această cotitură, iar ea este legată de limitele evoluției noastre.

Inteligența omenească este măsurată cel mai adesea prin intermediul unei metode dezvoltate în 1912 de psihologul german William Stern. Cunoașteți această metodă sub numele de „coeficient de inteligență“ sau IQ. Scorul se calculează împărțind rezultatul unui test de inteligență la vârsta exprimată în ani apoi

înmulțind rezultatul cu 100. Aproximativ 2,5% din populație obține un scor de peste 130, fiind considerați gânditori de elită, în vreme ce 2,5% au un scor sub 70, fiind considerați a avea dizabilități de învățare sau psihice. Chiar și cu câteva procentaje ca marjă de eroare, două treimi din populație se încadrează între 85 și 115 pe această scară. Cu toate acestea, suntem ceva mai inteligenți decât eram în trecut. De la începutul secolului XX, scorul mediu crește într-un ritm de aproximativ 3 puncte în fiecare deceniu, probabil datorită îmbunătățirii alimentației, educației și complexității mediului. [169] Ca urmare, nivelul mediu al oamenilor a început să crească. Dacă această tendință va continua, ne-am aștepta să avem mai multe genii până la finalul secolului decât avem acum. Între timp, evoluția noastră biologică se va intersecta cu inteligența artificială.

Pe măsură ce capacitatea noastră intelectuală se îmbunătățește, se va îmbunătăți și cea a inteligenței artificiale – însă nu putem folosi testele IQ pentru a măsura inteligența artificială. Folosim în schimb numărul de operații sau calcule pe secundă, cunoscute și sub numele de ops, pe care le poate efectua un calculator, pe care le putem compara apoi cu creierul uman. În funcție de sursă, numărul maxim de operații pe secundă pe care le pot efectua creierele noastre este de 1 exaflop, aproximativ un miliard de miliarde de operațiuni pe secundă, iar acestea reflectă multe activități care se petrec fără să fim conștienți de ele: micile mișcări pe care le facem când respirăm, procesarea vizuală continuă care se petrece când avem ochii deschiși și așa mai departe. Lansat în 2010, computerul chinezesc Tianhe-1 era cel mai rapid și mai puternic supercalculator, construit în întregime cu micro-

procesoare chinezești și care poate atinge un maxim teoretic de 1,2 petaflops. (Un petaflop înseamnă un milion de miliarde de operații pe secundă.) E rapid, dar nu se apropie de creierul omenesc. Apoi, în luna iunie 2018, IBM și Departamentul Energiei au lansat computerul Summit, care poate atinge 200 de petaflops, fiind construit în mod specific pentru inteligența artificială. [170] Această înseamnă că ne apropiem de o mașină cu o putere de calcul mai mare decât a noastră, chiar dacă ea nu ar putea încă trece testul Turing și să ne păcălească să credem că e un om.

Viteza nu este însă unicul parametru relevant. Dacă am accelera viteza creierului unui câine la 10^{15} ops, el nu ar putea dintr-o dată să se apuce să rezolve ecuații diferențiale- el doar ar alerga prin curte și ar mirosi și urmări mai multe obiecte. Creierul omului are o arhitectură mai complexă decât a unui câine: avem mai multe conexiuni între neuroni, proteine speciale și noduri cognitive complicate. [171] Chiar și așa, inteligența artificială poate fi extinsă de o manieră imposibilă pentru oameni fără a schimba arhitectura de bază a creierului nostru. Legea lui Moore, care susține că numărul de componente din circuitele integrate se dublează la fiecare doi ani pe măsură ce tranzistorii se micșorează continuă să se aplice și ne spune că evoluția computerelor urmează un tipar de creștere exponențială. Există din ce în ce mai multe date, precum și noi algoritmi, componente avansate sau noi modalități de a conecta rețelele neurale. Toate aceste evoluții generează mai multă putere. Spre deosebire de computere, noi nu putem modifica ușor structura creierelor noastre și arhitectura inteligenței omenești. Pentru a putea face asta ar trebui să (1) înțelegem în totalitate cum funcționează creierele noastre (2) să

modificăm arhitectura și substanțele chimice din creierele noastre iar aceste modificări să poată fi transmise generațiile următoare și (3) să așteptăm o perioadă lungă de timp până când vom face copii.

În ritmul actual, pentru ca IQ-ul oamenilor să crească cu 15 puncte e nevoie de 50 de ani. Iar pentru noi 15 puncte se observă imediat. Diferența dintre un creier din partea superioară a mediei, cu un IQ de 119, și unul cu un IQ de 134 „dotat" ar presupune o capacitate cognitivă semnificativ mai mare- efectuarea cu mai mare rapiditate a conexiunilor, stăpânirea mai ușoară a noilor concepte și o gândire mai eficientă. În același interval de timp, capacitățile cognitive ale inteligenței artificiale nu doar că ne vor depăși- ele vor deveni de nerecunoscut pentru noi pentru că nu avem din punct de vedere biologic suficientă putere de calcul pentru a înțelege ce este inteligența artificială. Pentru noi, o întâlnire cu o mașină super-inteligentă ar fi ca un cimpanzeu care ar participa la o ședință de consiliu local. Cimpanzeul ar putea să recunoască faptul că în cameră sunt oameni și că poate sta pe scaun, dar ce ar înțelege dintr-o discuție despre benzi de bicicletă într-o intersecție aglomerată? El nu dispune nici pe de parte de capacitatea cognitivă de a descifra limbajul folosit, darmite de argumentele și experiența care i-ar permite să combată propunerea benzilor de bicicletă. Pe termen lung în evoluția inteligenței și drumul către SAI, noi oamenii jucăm rolul cimpanzeului.

O inteligență artificială super-inteligentă nu e neapărat periculoasă, și nu subminează neapărat rolul pe care îl jucăm în cadrul civilizației. Cu toate aceste, SIA ar lua decizii de o manieră inconștientă folosind o logică de care am fi străini. Filosoful Nick

Bostrom de la Universitatea Oxford explică rezultatele posibile ale SIA folosind o parabolă cu agrafe de hârtie. Dacă am cere unei SAI să fabrice agrafe, ce s-ar petrece? Rezultatele fiecărei inteligențe artificiale, inclusiv cele care există deja, sunt stabilite prin valori și obiective. E posibil ca SIA să inventeze o nouă agrafă, mai bună, care să țină hârtiile chiar și când ar cădea pe jos. E posibil ca dacă nu suntem capabili să spunem câte agrafe ne *trebuie*, SIA să continue să fabrice la infinit agrafe, umplându-ne casele și birourile cu ele, dar și spitalele, școlile, râurile, lacurile, canalizarea și așa mai departe până când întreaga planetă e plină de agrafe. O SIA care acordă prioritate eficienței ar putea decide că oamenii stau în calea agrafelor, modificând întreaga planetă și transformând-o într-o fabrică de agrafe, provocând dispariția speciei noastre. [172] Iată de ce mulți experți în inteligență artificială, printre care mă număr și eu, sunt îngrijorați: dacă abilitățile cognitive ale SIA sunt de multe ori mai bune decât ale noastre (rețineți, nu suntem decât ceva mai inteligenți decât cimpanzeii) atunci suntem incapabili să ne imaginăm consecințele pe care le-ar putea avea aceste mașini inteligente asupra civilizației noastre.

Acesta este motivul pentru care cercetătorii din domeniul inteligenței artificiale folosesc frecvent cuvântul „explozie". Matematicianul și criptologul britanic I.J. Good a folosit prima dată termenul într-un eseu din 1965: „O mașină ultra-inteligentă ar putea crea mașini și mai performante; s-ar produce indenegabil o „explozie a inteligenței" iar inteligența omenească ar rămâne cu mult în urmă. Astfel, prima mașină ultra-inteligentă ar fi ultima invenție pe care omul ar trebui să o mai facă, atât timp cât mașina e suficient de docilă încât să ne spună cum să o ținem sub control.[173]

Cei Nouă Titani construiesc structuri și sisteme care -ei speră- vor încuraja într-o bună zi o astfel de explozie, creând un teren propice unor soluții, strategii, concepte, structuri și abordări complet noi, la care nici măcar cei mai inteligenți dintre informaticieni nu s-au gândit. Astfel ar urma evoluții din ce în ce mai rapide, oportunități și creșterea cifrei de afaceri. În termeni de specialitate, o astfel de explozie se numește „auto-îmbunătățire recursivă", referindu-se la un ciclu prin care inteligența artificială devine prin propriile forțe mai rapidă, mai bună și mai inteligentă prin modificarea capacităților sale. Astfel, inteligența artificială ar putea să preia controlul și să-și făurească propriul destin. Îmbunătățirile s-ar putea produce la fiecare oră, sau chiar instantaneu.

„Explozia inteligenței" care va urma descrie nu doar viteza super-computerelor sau puterea algoritmilor ci și răspândirea la scară largă a mașinilor gânditoare inteligente care urmăresc auto-îmbunătățirea recursivă. Imaginați-vă o lume în care sisteme mult mai avansate decât AlphaGo Zero și NASNet nu doar că iau decizii strategice în mod autonom, ci și colaborează și concurează ca parte a unei comunități globale. O lume în care li se cere să evolueze, în principal pentru a ajuta oamenii- scriind noi generații de programe, evoluând și auto-îmbunătățindu-se, însă la o viteză uluitoare. Inteligențele artificiale ar crea noi agenți, programându-i cu un scop și un set de sarcini, iar ciclul s-ar repeta în continuu, de mii de miliarde de ori, provocând atât schimbări minore cât și colosale. Singura epocă istorică în care s-a produs un cataclism evolutiv de o magnitudine similară a avut loc aproximativ acum 542 de milioane de ani în epoca cambriană, când diversificarea ra-

pidă a formelor de viață a condus la apariția unor noi ființe complexe și a transformat pentru totdeauna planeta noastră. Fostul director de programe de la DARPA Gill Pratt crede că ne aflăm în mijlocul unei explozii de tip cambrian- o perioadă în care inteligența artificială învață din experiența tuturor inteligențelor artificiale, după care viața pe Pământ ar putea fi complet diferită față de prezent.[174]

Acesta e motivul pentru care cei Nouă Titani, investitorii și acționarii lor, agențiile guvernamentale și politicienii, cercetătorii trebuie să recunoască semnalele de alarmă și să se gândească mai critic nu doar la viitorul IAÎ care este creată acum, dar și la IAG și SIA care se întrevăd la orizont. Și voi jucați un rol la fel de important în cadrul acestui proces. Evoluția inteligenței constituie un spectru, pe care coexistă atât oamenii cât și mașinile. Valorile celor Nouă Titani sunt deja profund întipărite în algoritmii, sistemele și structurile existente. Aceste valori vor fi transmise mai departe către milioane de noi generații de inteligențe artificiale care evoluează, și în curând unor mașini gânditoare capabile de inteligență generală.

Tranziția de la IAÎ la SIA va dura probabil următorii 70 de ani. În momentul de față sunt dificil de definit etapele concrete pentru că progresul în domeniul inteligenței artificiale depinde de mai mulți factori și oameni: noii membri incluși în triburile inteligenței artificiale, deciziile strategice luate de cei Nouă Titani, disputele comerciale și geopolitice, ca să nu mai vorbim de evenimentele haotice sau de hazard. În propriile mele modele, consider că IAG va apărea în jurul anului 2040. Pare că e un viitor îndepărtat, așa că voi încerca să creez un context pentru această

dată. Vom fi avut trei sau patru preşedinţi diferiţi la Casa Albă. (În absența unor probleme de sănătate, preşedintele Xi Jinping va fi în continuare la putere.) Voi avea 65 de ani când sistemele IAG vor putea efectua propria cercetare în domeniul inteligenței artificiale. Copilul meu de clasa a doua va avea 30 de ani, iar în acel moment ar putea să citească un bestseller New York Times scris în întregime de o maşină. Tata va avea peste 90 de ani iar toţi specialiştii (cardiologi, nefrologi, radiologi) vor fi IAG, dirijate şi controlate de un medic generalist hiper-specializat, care va fi atât un medic cât şi un analist de date. Apariția SIA ar putea urma la scurt sau mult timp după aceea, undeva între 2040 şi 2060. Nu e obligatoriu ca până în 2070 SIA să fi distrus viața pe Pământ sub miliarde de agrafe. Cu toate acestea, nici această variantă nu poate fi exclusă.

Poveştile pe care trebuie să ni le spunem

Pentru a putea planifica viitorul inteligenței artificiale avem nevoie de noi poveşti folosind date din lumea reală. Dacă suntem de acord că inteligența artificială va evolua pe parcurs, va trebui să construim scenarii care să descrie intersectarea celor Nouă Titani cu forțele economice şi politice care îi ghidează, precum şi rolul omenirii pe măsură ce inteligența artificială trece de la etapa aplicațiilor înguste la inteligență generală şi în cele din urmă la super-inteligență.

Pentru că viitorul nu s-a produs încă, nu putem şti cu siguranță toate rezultatele posibile ale acțiunilor noastre din prezent. Din acest motiv, scenariile următoare sunt scrise folosind diferite

încadrări afective pentru următorii cincizeci de ani. Primul este scenariul optimist, în care ne întrebăm ce se întâmplă dacă cei Nouă Titani decid să adopte schimbări radicale pentru a se asigura că inteligența artificială va fi în beneficiul tuturor. Am o precizare importantă de făcut: scenariile „optimiste" nu sunt neapărat încurajatoare sau motivante. Ele nu conduc întotdeauna la o utopie. Într-un scenariu optimist presupunem că se iau cele mai bune decizii și că se îndepărtează obstacolele care stau în calea reușitei. În ce ne privește, aceasta ar presupune ca cei Nouă Titani să-și schimbe abordarea privind inteligența artificială, iar pentru că ei iau tot timpul cele mai bune decizii, viitorul nostru va fi mult mai bun decât prezentul. E un scenariu în care mi-ar plăcea să trăiesc, și un viitor pe care-l putem avea dacă lucrăm împreună.

Următorul este un scenariu pragmatic care descrie cum ar arăta viitorul dacă cei Nouă Titani nu fac pe termen scurt decât mici îmbunătățiri. Presupunem că deși toate persoanele interesate recunosc că inteligența artificială nu se îndreaptă într-o direcție bună, nu va exista o colaborare pentru a pune bazele unor schimbări semnificative și de durată. Câteva universități vor introduce cursuri obligatorii de etică, Mafia G formează parteneriate la nivel de industrie pentru a combate riscurile dar fără să-și modifice cultura organizațională a membrilor; politicienii se concentrează asupra alegerilor și pierd din vedere planurile ambițioase ale Chinei. Scenariul pragmatic nu este unul al marilor schimbări-el ține cont de fluxul și refluxul demersurilor de îmbunătățire ale omenirii. Acest scenariu ține cont de faptul că liderii din mediul politic și de afaceri sunt cât se poate de dispuși să saboteze viitorul de dragul unor câștiguri imediate, pe termen scurt.

În ultimul rând, scenariul catastrofal explică ce s-ar întâmpla dacă toate semnalele de alarmă ar fi ignorate și nu ne-am construi un plan pentru viitor, iar cei Nouă Titani vor continua să concureze unii împotriva celorlalți. Dacă alegem să mergem mai departe pe aceleași baze, încotro am putea ajunge? Ce s-ar întâmpla dacă inteligența artificială și-ar continua parcursul actual în Statele Unite și China? Crearea unei schimbări sistemice -care ar putea împiedica scenariul catastrofal- e o muncă grea, de durată, și fără un final concret. Acesta este motivul pentru care scenariul catastrofal pare atât de înfricoșător și amănuntele atât de tulburătoare, fiindcă în prezent pare că ne îndreptăm către acest scenariu.

Am studiat, modelat și conceput aceste situații pentru a descrie rezultatele începând cu anul 2029. La baza scenariilor se regăsesc câteva teme principale, printre care se numără: mobilitatea și oportunitățile economice, productivitatea forței de muncă, îmbunătățirea structurilor sociale, dinamica de putere a celor Nouă Titani, relațiile dintre Statele Unite și China și expansiunea sau retragerea globală a comunismului și democrației. Voi arăta cum valorile noastre sociale și culturale s-ar putea schimba pe măsură ce inteligența artificială ajunge la maturitate: felul în care definim creativitatea, felul în care relaționăm unii cu ceilalți și ce credem despre viață și moarte. Deoarece scenariile urmăresc să ne ajute să înțelegem cum ar arăta viața noastră în timpul tranziției de la IAÎ la SIA, am inclus exemple legate de casele, locurile de muncă, educație, sănătate, ordine publică, urbanism, infrastructură locală, politică și siguranță națională.

Unul dintre rezultatele probabile ale inteligenței artificiale și un element comun al tuturor celor trei scenarii îl reprezintă

apariția unui „Dosar de Date Personale", sau DDP. Acesta constituie un registru unic care ar cuprinde toate datele generate de activitatea noastră digitală (gândiți-vă la internet sau la telefoane mobile) dar ar include și date din alte surse: notele de la școală, istoricul de muncă (diplome, foști și actuali angajatori); date cu caracter juridic (căsătorii, divorțuri, arestări); date financiare (ipoteci, scoruri de credit, împrumuturi, impozite); date despre călătorii (țări vizitate, vize); istoricul relațiilor (aplicații online); date de sănătate (arhive electronice de sănătate, analize genetice, activitatea sportivă) și istoricul de cumpărături (magazine online, folosirea cupoanelor de reduceri). În China un astfel de DDP ar include și datele privind scorul de credit social amintit în capitolul anterior. Inteligența artificială creată de către cei Nouă Titani ar putea atât să învețe despre voi pe baza acestui DDP cât și să îl folosească pentru a lua decizii automate și a vă oferi o gamă de servicii. DDP ar putea fi moștenibil -un istoric cuprinzător transmis copiilor pentru a-l putea folosi- și ar putea fi gestionat temporar sau proprietatea exclusivă a unei companii din cadrul celor Nouă Titani. DDP joacă un rol important în scenariile pe care urmează să le citiți.

Aceste DDP nu există încă, însă din punctul meu de vedere există deja semnale care par să arate că ne îndreptăm către un viitor în care multiplele surse de date personale vor fi adunate laolaltă într-un singur fișier pus la dispoziție și actualizat de către cei Nouă Titani. De fapt, luați deja parte la acest sistem și folosiți o variantă prototip de DDP. E vorba de adresa voastră de e-mail.

Adresa de e-mail a unei persoane obișnuite a fost făcută să acționeze și ca o modalitate de autentificare; numărul de telefon

este folosit pentru validarea tranzacțiilor iar telefonul este folosit pentru a-i găsi în lumea fizică. Dacă folosiți Gmail, Google -și prin extensie, și sistemele lor de inteligență artificială, această companie știe mai multe despre voi decât știu partenerii voștri. Sistemul cunoaște numele și adresele de e-mail ale tuturor persoanelor cu care corespondați, precum și informații demografice legate de ele (vârstă, gen, localitate de domiciliu). Google știe când aveți tendința să vă verificați e-mailul și în ce împrejurări. Pe baza e-mailului știe unde călătoriți, de ce sume de bani dispuneți și ce cumpărați. Dacă faceți poze cu un telefon Android, acesta recunoaște fețele prietenilor și rudelor, putând detecta abaterile de la tipar pentru a trage concluzii: de exemplu multe poze noi cu aceeași persoană ar putea însemna că aveți o nouă prietenă (sau o relație extraconjugală). Google știe de toate întâlnirile, programările la doctor și planurile voastre de a vă duce la sală. Google știe dacă țineți Ramadanul, Rosh Hashanah, dacă vă duceți la biserică sau dacă nu vă interesează nicio religie. Google știe unde ar trebui să fiți de regulă marți după-amiază, chiar dacă sunteți în altă parte. Sistemul știe ce căutați, folosind degetele sau vocea, așa că-și poate da seama dacă aveți prima dată un avort spontan, dacă învățați cum să gătiți paella, dacă aveți probleme de identitate sexuală, dacă vă trece prin cap să renunțați la carne sau să vă căutați o nouă slujbă. Aceste date sunt interconectate, iar sistemul învață de pe urma lor, monetizând și transformând în produse concluziile în timp ce vă sugestionează către anumite direcții prestabilite.

În prezent, Google are acces la aceste informații în mod voluntar doar pe baza unui singur produs- adresa de Gmail- pe care probabil o folosiți și pentru a cumpăra produse de pe Amazon sau

a vă autentifica în Facebook. Nu mă plâng, ci doar descriu viața modernă. Pe măsură ce inteligența artificială progresează, un DDP mai explicit va permite celor Nouă Titani o mai mare eficiență, astfel încât ne vor sugestiona să acceptăm și să adoptăm aceste DDP, chiar dacă nu înțelegem pe deplin ce presupune folosirea lor. Desigur că în China aceste DDP au fost deja introduse ca proiecte pilot în cadrul sistemului de credite sociale.

„Ne spunem povești pentru a putea trăi," scria Joan Didion în cartea sa *The White Album*. „Interpretăm ceea ce vedem, și alegem cele mai fezabile dintre variante." Trebuie să avem de ales când vine vorba de inteligența artificială. A venit vremea să folosim informația disponibilă pentru a ne spune povești-scenarii care să descrie cum am putea convieți cu mașinile gânditoare.

CAPITOLUL CINCI

PROSPERITATE ÎN A TREIA EPOCĂ A INFORMATICII: SCENARIUL OPTIMIST

Suntem în anul 2023 și am luat cele mai bune decizii în privința inteligenței artificiale- am modificat parcursul de dezvoltare al inteligenței artificiale, construim viitorul și vedem deja schimbări permanente în bine. Triburile inteligenței artificiale, universitățile, cei Nouă Titani, agențiile guvernamentale, investitorii, cercetătorii și oamenii obișnuiți au ținut cont de semnalele timpurii de alarmă.

Înțelegem că există o singură schimbare care va rezolva problemele deja create și acum cea mai bună strategie este de a ne ajusta așteptările legate de viitorul inteligenței artificiale. Recunoaștem că inteligența artificială nu este doar un produs al Silicon Valley- un bun care să fie valorificat cât timp piața este în creștere.

Înainte de orice, recunoaștem de ce China a investit strategic în inteligența artificială și cum parcursul de dezvoltare al

inteligenței artificiale se potrivește cu rolul pe care China și-l asumă în lumea viitorului. China nu încearcă să modifice balanța comercială ci să obțină un avantaj absolut asupra Statelor Unite din punct de vedere al puterii economice, dezvoltării forței de muncă, influenței geopolitice, forței militare, influenței sociale și al politicilor de mediu. Dându-și seama de acest lucru, politicienii noștri au construit, cu sprijinul deplin al Mafiei G și al triburilor inteligenței artificiale, o coaliție internațională care să protejeze și să păstreze inteligența artificială ca un bun public. Această coaliție presează China și folosește pârghiile economice pentru a contracara folosirea inteligenței artificiale cu scopul de a supraveghea cetățenii și a consolida puterea regimului comunist.

Recunoscând că China folosește inteligența artificială pentru a-și atinge obiectivele economice și militare în timp ce răspândește idealurile comunismului și-și consolidează puterea asupra societății, guvernul federal al Statelor Unite alocă fonduri substanțiale pentru a sprijini dezvoltarea inteligenței artificiale, ceea ce reduce presiunea asupra Mafiei G de a căuta profituri rapide. Folosind cursa spațială a anilor '50 drept precedent, devine evident cât de ușor poate fi depășită America de alte țări în absența unei coordonări la nivel național. Devine de asemenea cât se poate de evident ce fel de influență poate să exercite America în domeniul științei și tehnologiei atunci când avem o strategie națională coordonată- grație guvernului federal au apărut tehnologii precum GPS sau internetul.

Nici inteligența artificială și nici finanțarea nu sunt politizate și toată lumea este de acord că reglementarea Mafiei G sau inteligenței artificiale reprezintă o abordare greșită. O reglemen-

tare strictă, dură ar deveni desuetă chiar din momentul în care ar fi adoptată, ar sufoca inovația și ar fi dificil de pus în practică. Cu sprijinul ambelor partide politice, americanii sprijină sporirea cheltuielilor federale în domeniul inteligenței artificiale folosind strategia chineză drept model. Se alocă fonduri pentru cercetare și dezvoltare, studii de impact asupra economiei și forței de muncă, studii de impact social, programe destinate diversității, inițiative medicale și de sănătate publică și infrastructură și pentru revitalizarea sistemului public de educație, oferind salarii atractive profesorilor și o programă care oferă tuturor o mai bună pregătire pentru un viitor din ce în ce mai automatizat. Nu mai presupunem că Mafia G poate să-și servească în egală măsură stăpânii politici și financiari, sau că piețele libere și spiritul nostru antreprenorial vor produce cele mai bune rezultate pentru inteligența artificială și pentru omenire.

Având la dispoziție o strategie națională și finanțare, nou înființata coaliție a Mafiei G semnează acorduri oficiale multilaterale pentru a colabora la dezvoltarea viitorului inteligenței artificiale. Coaliția Mafiei G definește și adoptă standarde care acordă prioritate dincolo de orice unui parcurs de dezvoltare pentru inteligența artificială care ține cont de interesele democrației și societății. Coaliția acceptă unirea tehnologiilor din domeniul inteligenței artificiale. Prin colaborare rezultă chipuri, structuri și arhitecturi de rețea superioare, care înlocuiesc sistemele rivale de inteligență artificială și o comunitate de programatori cu mai multe obiective. Această colaborare permite și cercetătorilor să

caute noi oportunități, iar astfel toată lumea câștigă.

Coaliția Mafiei G adoptă transparența ca valoare fundamentală, rescriind în mod radical reglementările, condițiile și fluxurile de lucru pentru a acorda prioritate înțelegerii și educației. Acest proces se produce voluntar, evitând astfel intervenția autorităților. Seturile de date, algoritmii de pregătire și structurile de rețele neurale sunt transparentizate de o manieră care să protejeze doar acele secrete comerciale și informații confidențiale care ar putea provoca pagube economice unuia dintre membrii coaliției. Departamentele juridice ale Mafiei G nu mai petrec ani de zile căutând sau dezbătând lacune ale legislației și nu mai caută să amâne adoptarea măsurilor legate de transparență.

Știind că automatizarea urmează să se producă, Mafia G ne ajută să creăm diferite scenarii pentru problema șomajului și să ne pregătim forța de muncă pentru cea de-a treia eră a informaticii. Cu ajutorul acestora, nu ne temem de inteligența artificială, ci o vedem ca pe o ocazie extraordinară pentru a spori creșterea economică și prosperitatea fiecăruia dintre noi. Conducătorii Mafiei G reușesc să învingă zvonurile și să ne prezinte abordări mai eficiente pentru a ne pregăti și educa pentru slujbele viitorului.

Strategia națională a Statelor Unite și formarea Coaliției Mafiei G inspiră și liderii altor țări democratice din lume să sprijine dezvoltarea globală a inteligenței artificiale pentru binele tuturor. Universitatea Dartmouth găzduiește ediția inaugurală a forumului inter-guvernamental, într-o adunare care se aseamănă cu cea din vara anului 1956, unde se întâlnesc lideri din economiile cele mai

avansate ale lumii: secretari, miniștri, prim-miniștri și președinți din Statele Unite, Regatul Unit, Japonia, Franța, Canada, Germania, Italia și alte țări din Uniunea Europeană, precum și cercetători din domeniul inteligenței artificiale, sociologi, economiști, futurologi, specialiști în teoria jocului, politică și alții. Spre deosebire de grupul omogen de participanți la primul atelier de la Dartmouth- bărbați cu origini asemănătoare, de data aceasta experții și liderii includ un spectru larg de oameni și perspective asupra lumii. Aflându-se pe același loc sfânt unde s-a născut inteligența artificială modernă, acești lideri acceptă să mijlocească și să coopereze în cadrul unor inițiative și politici comune în domeniul inteligenței artificiale. Inspirați de mitologia greacă și de mama Pământului, ei formează GAIA: Alianța Globală de Augmentare a Inteligenței.

Exclusă din GAIA, China constată că influența sa globală este în scădere. Colaborarea internațională nu are un impact financiar negativ asupra membrilor chinezi ai celor Nouă Titani-Baidu, Tencent și Alibaba- care continuă să ofere numeroase servicii cetățenilor chinezi. Cu toate acestea, multe dintre planurile pe termen lung ale Chinei- printre care și inițiativa Noului Drum al Mătăsii- sunt afectate, deoarece partenerii ies din etapa pilot și recrutarea unor noi aliați se dovedește dificilă.

Nu sugerez că toate problemele inteligenței artificiale dispar peste noapte. Comunitatea inteligenței artificiale anticipează și se așteaptă ca inteligența artificială îngustă să continue să facă greșeli din cauza perspectivelor limitate asupra lumii ale membrilor inițiali ai tribului inteligenței artificiale. Acceptăm că prejudecățile politice, de gen, rasă și statut social nu vor dispărea imediat. Țările

membre GAIA semnează acorduri prin care se obligă în mod explicit să valorizeze siguranța în defavoarea vitezei și să aloce resurse considerabile în vederea curățării tuturor sistemelor existente: bazele de date și algoritmii deja implementați, structurile pe care acestea se sprijină, produsele pentru companii care înglobează inteligența artificială (precum cele folosite de bănci sau de forțele de ordine) și produsele pentru consumatori care folosesc inteligența artificială pentru sarcinile zilnice (difuzoarele, ceasurile și telefoanele noastre inteligente). GAIA pretinde -și recompensează- răspunderea în fața publicului.

În cadrul GAIA, se ia decizia ca Dosarele de Date Personale (DDP) să fie prelucrate la fel cu registrele distribuite de blockchainuri. Registrele distribuite folosesc mii de calculatoare independente pentru a înregistra, partaja și sincroniza tranzacțiile. Ele sunt structurate pentru a nu păstra datele centralizat, sub umbrela unei singure companii sau agenții. Din cauză că Coaliția Mafiei G adoptă un set de standarde și oferă tehnologii de inteligență artificială unitare, DDP-urile noastre nu au nevoie de o companie coordonată central care să gestioneze tranzacțiile. Ca urmare, indivizii sunt deținătorii propriilor DDP-uri, care sunt pe cât de publice sau de private ni le dorim și sunt complet interoperabile-le putem conecta la unul sau toate dintre serviciile Mafiei G simultan, dar și în alte domenii, cum ar fi la doctor, la școală sau la infrastructura orașului nostru. Membrii Mafiei G sunt custozii inteligenței artificiale și ai datelor noastre, însă nu și proprietarii lor. DDP-urile sunt moștenibile: putem transmite datele noastre copiilor având capacitatea de a acorda permisiuni (de vizibilitatea completă, limitată sau zero) asupra diferitelor aspecte ale datelor proprii.

Pe măsură ce inteligența artificială se maturizează și trece de la etapa aplicațiilor înguste la mașinile gânditoare capabile de inteligență generală, triburile inteligenței artificiale și Mafia G ne-au câștigat încrederea. Nu sunt doar companii care produc aplicații utile – Google, Microsoft, Apple, Facebook, IBM și Amazon sunt la fel de importante pentru America și valorile americane ca și baseball-ul, exprimarea liberă și ziua de 4 iulie. Comunismul este marginalizat. Țările care respectă dreptul cetățenilor de a se exprima și de a deține proprietăți, care sprijină libertatea religioasă, care sprijină oamenii de toate genurile, etniile, rasele și orientările sexuale, care sunt de acord că guvernul trebuie să fie în slujba cetățenilor, care guvernează prin intermediul unor reprezentanți aleși și care pun în balanță atât libertatea individuală și siguranța publică sunt aliate și conlucrează pentru viitorul inteligenței artificiale și omenirii.

2029: Sugestionarea confortabilă

Grație colaborării între membrii Mafiei G și înființării GAIA, se semnează numeroase acorduri comerciale, cetățenii lumii au acces mai bun și mai ieftin la produse și servicii bazate pe IAÎ. GAIA se întrunește în mod regulat, publicând toate dezbaterile, în timp ce grupurile sale multinaționale de lucru țin pasul cu progresele tehnologice.

Casele clasei de mijloc se bazează pe inteligența artificială pentru a-și ușura viața. Aparatele, platformele și alte servicii sunt

interoperabile chiar și între țări, deși în urmă cu câteva decenii licențele și restricțiile asupra datelor împiedicau accesul internațional. Mașinile de spălat și uscătoarele folosesc mai puțină energie, sunt mai eficiente și se conectează la sistemele orașelor inteligente pentru a partaja date. Cu acordul nostru, permitem ca ciclurile de spălare să se efectueze atunci când solicită cel mai puțin rețelele de apă și electricitate.

IAÎ sprijină calculele senzoriale, ceea ce înseamnă că putem strânge și interoga lumea reală folosind date senzoriale: văzul, mirosul, auzul, gustul și atingerea. Folosim scannere ținute în mână, cu camere inteligente și vizualizare digitală în bucătărie. Spectrometrul inclus într-o baghetă cu IAÎ din bucătărie citește lumina reflectată dintr-un avocado pentru a vă spune că nu se va coace până în weekend-în timp ce uleiul de măsline pe care l-ați cumpărat de la reduceri nu e pur ci un amestec de trei uleiuri diferite. Un alt senzor din bucătărie a detectat că puiul pus la rotisor urmează să se ardă. La etaj, un senzor tactil vă alertează că bebelușul a scăpat (din nou) din pătuț.

Mafia G a încheiat parteneriate cu alte companii din domeniul realității augmentate, îmbunătățind semnificativ viața persoanelor care suferă de Alzheimer și demență. Ochelarii inteligenți recunosc automat oamenii, obiectele și locurile, ajutându-ne rudele să-și aducă aminte și să aibă o viață mai împlinită.

Inițial, credeam că produsele și serviciile Mafiei G ar provoca izolare socială- că vom sta cu toții singuri în case, interacționând prin avataruri digitale în timp ce ne pierdem complet legătura cu

lumea din afară. Ne-am înșelat complet. Platformele și aparatura Mafiei G ne-au oferit noi moduri de a socializa pe viu. Petrecem mai mult timp în cinemauri cu realitate mixtă, care oferă un divertisment multi-senzorial. Există peste tot săli de jocuri cu realitate mixtă. E din nou ca în anii '80, doar că s-a schimbat ceva: jocurile, experiențele și sălile de întâlnire cu realitate mixtă sunt ieftine, și ele pot fi folosite și de persoanele cu tulburări de vedere și auz. Ne ducem la petreceri liniștite, unde purtăm căști colorate fără fir și ascultăm toată noaptea mixurile DJ-ului preferat. Acum cu toții pot dansa împreună, într-o experiență comună, chiar dacă își detestă reciproc gusturile muzicale. Grație Mafiei G suntem mai conectați unii la ceilalți-și la lumea reală- decât ne-am fi putut vreodată imagina.

Pentru casele mai bogate, aplicațiile IAÎ oferă chiar și mai multe dotări. În grădină, senzorii măsoară constant nivelul de umiditate și compară datele cu prognozele microclimatului. Sistemele simple de irigații udă plantele automat, dar doar dacă este necesar. Inteligența artificială poate calcula nivelul optim de hidratare, ceea ce înseamnă că s-a zis cu aparatele care măsoară când trebuie udate plantele, dar și cu begoniile moarte.

În interiorul acestor case mai bogate, sistemul Akira de la Amazon (cu o voce care nu sună nici a bărbat, nici a femeie) operează în mai multe limbi, indiferent de accent, și poate comunica cu ușurință cu ochelarii inteligenți de la Apple dar și datele personale gestionate de Google. Mașinile de spălat și uscătoarele sunt prevăzute cu mici drone cu brațe și o nouă funcționalitate denumită „modul Kondo", în onoarea lui Marie Kondo, expertul japonez în ordonarea casei. Rufele sunt spălate și uscate în funcție de

ciclul de cerere și ofertă al grilei de alimentare cu curent electric, iar apoi hainele sunt preluate de o mică dronă care le împachetează și le sortează după culori.

În Statele Unite, cumpărăturile și livrarea lor se efectuează complet automat. Nu rămâneți niciodată fără tampoane sau pastă de dinți. Inteligența artificială stă la baza unor sistem predictive de achiziții care își dau seama înaintea voastră, pe baza achizițiilor anterioare și a DDP-urilor voastre când trebuie să refacă stocurile. Prin intermediul Amazon puteți cumpăra legume și carne proaspătă din zona voastră, precum și toate articolele de bază, cum ar fi cerealele pentru micul dejun, hârtia igienică sau chipsurile de cartofi. Serviciile de livrat truse de ingrediente la domiciliu, lansate în urmă cu un deceniu de Blue Apron și HelloFresh sunt legate de PDP-ul unei case. În schimbul unei sume mai mari de bani în fiecare săptămână, cumpărăturile voastre vor include ingrediente pentru toate rețetele favorite, dar și trei noi mese-rețete aliniate automat cu preferințele, antipatiile, alergiile și nevoile nutritive ale fiecărui membru al familiei.

Veți continua desigur să cumpărați și în lumea reală, dar veți putea alege să vă lăsați portofelul acasă. Tehnologia care stă la baza sistemelor comerciale Amazon Go a devenit coloana vertebrală a magazinelor cu servire rapidă în care o bună parte a inventarului este deja la vedere sau poate fi ușor înlocuit. Camerele inteligente supraveghează constant cumpărătorii, recunoscând amprentele faciale și conținutul coșurilor de cumpărături. Putem cheltui până la 100 de dolari fără să fim nevoiți să interacționăm cu un angajat. În magazinele cu suprafețe mai mari (de ex., magazine universale, de

mobilă sau de bricolaj) sau cele care vând marfă mai scumpă (de ex., bijuterii, genți și electronice) avem opțiunea să plătim cu fața.

Unii copii se joacă cu animale vii, în timp ce familiile cu un stil de viață mai agitat optează pentru animale robotizate. Mici câini și pisici -niște receptacule drăguțe ale inteligenței artificiale- folosesc calculele senzoriale și învățarea profundă în timp ce sunt dresați să respecte regulile casei. Dispunând de camere avansate în fosele oculare, de blană cu senzori tactili și de capacitatea de a recunoaște modificările subtile ale vocilor noastre, animalele robotizate sunt mult mai empatice decât cele vii, chiar dacă sunt mai puțin calde și pufoase.

Oricine, indiferent de nivelul de venit, se bucură de faptul că i se sugestionează un stil de viață mai sănătos. Mafia G ne aduce aminte în timpul zilei să facem alegeri mai sănătoase. În timp ce vă îndreptați spre muncă și așteptați liftul, ceasul vibrează un pic pentru a vă face să vă uitați în jos: vă arată o hartă simplă a clădirii de birouri, iar o săgeată vă îndreaptă către scări. Cu siguranță că puteți dezactiva această opțiune, însă majoritatea oamenilor preferă să o lase pornită. Și antrenamentele personale sunt optimizate. Folosind dosarul vostru de date personale, istoricul medical și datele senzoriale strânse din multe alte surse- căștile fără fir prin care ascultați muzică, țesătura inteligentă din care este croit sutienul sportiv- echipamentul de fitness vă ghidează prin exerciții personalizate. După ce ați terminat antrenamentul, senzorii vă ajută să vă răcoriți, monitorizându-vă metabolismul și inima. Datorită Mafiei G, comunitățile noastre sunt mai sănătoase, iar viețile noastre sunt mai lungi.

Prin faptul că Mafia G a agreat un singur standard pentru DDP-uri, interfețele utilizatorilor, formatul de arhivare a datelor, protocoalele și structurile folosite în sistemul medical sunt standardizate. Ca urmare, sistemul de sănătate este mult mai eficient. Politicienii s-au certat preț de decenii referitor la asistența sanitară în America, iar insistența Mafiei G pe algoritmi și date standardizate în sănătate s-a dovedit a fi cel mai bun leac.

Indiferent ce doctor vede un pacient, sau la ce spital este internat, datele acestuia pot fi accesate de orice persoană responsabilă cu îngrijirea lui. Ele sunt disponibile și pentru orice persoană căreia pacientul i-a acordat permisiunea. Datele din majoritatea testelor de laborator, analizelor și scanărilor sunt analizate de inteligența artificială și nu de oameni, conducând la o precizie mai mare și la procesarea mai rapidă a rezultatelor. Sistemul IBM poate detecta anomalii celulare pentru a depista semnele timpurii ale cancerului- precum și care sunt celulele afectate din corp. Sistemul Google ajută doctorii să prevadă rezultatul probabil al diferitelor medicamente și tratamente, precum și să estimeze când va muri un pacient, ajutând personalul medical să ia decizii mai bune în privința tratamentului pentru fiecare pacient. Aflat în spital, un pacient primește medicamentele acasă ca urmare a sincronizării între farmacia Amazon și DDP. Chiar și dacă în istoricul medical al pacientului sunt pagini întregi de rețete scrise de mână de către doctori- și chiar dacă aceste observații clinice nu sunt foarte amănunțite- sistemul de vizualizare digitală și analiză a tiparelor al Mafiei G completează aceste detalii, transformând aceste informații în date structurate, care pot fi folosite și extrase doar pentru pacient, sau anonimizate și combinate cu alte date despre

pacienți pentru a sprijini comunitatea medicală (formată atât din oameni cât și din inteligențe artificiale) în a-și spori cunoașterea și experiența.

Diagnosticul, tratamentul și îngrijirea nu sunt oferite doar în spitale fizice, ceea ce înseamnă că mult mai mulți oameni din Statele Unite au acum un acces mai bun la sănătate. Deși relativ recente, unii furnizori oferă acum servicii de medicină la distanță. Toaletele TOTO, dotate cu spectrofotometru și eprubete pot folosi detectarea tiparelor pentru a diagnostica un nivel crescut sau scăzut de glucoză sau proteine, precum și bacteriile sau hematiile. După câteva secunde, DDP reflectă o posibilă infecție urinară sau semnele incipiente ale pietrelor la rinichi. Tratamentele simple- cum ar fi antibioticele pentru infecții- sunt comparate cu DDP-ul vostru, trimise medicului de familie, iar dacă acesta aprobă, sunt livrate automat la voi acasă, la muncă sau în oraș la cină.

Periuțele de dinți, dotate cu mici senzori de salivă analizează lichidul pentru a vedea care este starea generală de sănătate. Cu fiecare periaj, inteligența artificială monitorizează nivelul hormonilor, electroliților și anticorpilor, verificând dacă acesta este constant de-a lungul timpului. Mafia G a schimbat standardele de îngrijire: testele de diagnosticare elementară nu sunt doar pentru pacienții bolnavi, ci fac parte din menținerea unui stil de viață sănătos. Această evoluție a schimbat la rândul ei caracterul medicinei, ea trecând de la o abordare reactivă la una predictivă, profilactică.

Și alte aspecte ale vieții cotidiene- printre care și viața amoroasă sau sexul- sunt mai bune grație inteligenței artificiale. Algoritmii evolutivi s-au dovedit o soluție mai inteligentă pentru

persoanele care-și caută dragostea pe internet decât aplicațiile și paginile de internet existente până atunci. Cercetătorii au stabilit că oamenii sunt pur și simplu prea complecși pentru a putea fi reduși la un set de date pe care un singur algoritm trebuie să le compatibilizeze. Pe lângă acest aspect, avem tendința să completăm profilurile online oferind informații care descriu mai degrabă cum ne-am dori să fim -și nu cum suntem de fapt. Algoritmii evolutivi preiau datele din DDP-uri și ne verifică toate profilurile din bazele de date. Ne alegem un obiectiv – de la „vreau doar să mă distrez" la „sunt gata să mă căsătoresc"- apoi adăugăm parametrii de limitare (trebuie să fie de religie mozaică, și să trăiască la maximum 70 de km de Cleveland), iar algoritmul evolutiv generează o listă de oameni cu care avem cele mai mari șanse să ne atingem acel obiectiv. Dacă ne dorim, sistemul va consulta agenda și preferințele de activități și va programa automat data și locul întâlnirii. După câteva întâlniri (sau dacă prima întâlnire nu a mers prea bine), ne-ar putea interesa să folosim un algoritm generativ care să creeze conținut pornografic personalizat. În funcție de preferințele noastre, inteligența artificială creează scene care să ne excite, să ne inspire sau să ne învețe, cu personaje ale căror voci, corpuri și stiluri se mulează pe dorințele noastre.

Grație Mafiei G, inteligența artificială nu ne lasă impresia că ar înlocui creativitatea omenească, ci că o completează- e un instrument prin care să ne mărim și sporim inteligența. În cadrul firmelor de arhitectură, inteligența artificială generează mii de posibile clădiri pe baza modelelor preferate de client și a constrân-

gerilor precum și alege și ierarhizează planurile câștigătoare pe baza predicțiilor privind fezabilitatea luând în calcul termenul de execuție, materialele disponibile și bugetul; dificultatea obținerii autorizațiilor necesare și dacă clădirea afectează negativ traficul pietonal. Investitorii imobiliari folosesc inteligența artificială pentru a simula durabilitatea pe termen lung, dat fiind climatul unei zone dar și alți factori de mediu. Meseriașii calificați -tâmplari, electricieni și instalatori-folosesc ochelari cu realitate mixtă oferiți de Google, Microsoft și o companie pe nume Magic Leap pentru a vedea prin pereți, a verifica dacă proiectul se potrivește cu execuția și a detecta de la bun început eventualele probleme.

Folosirea creativă a inteligenței artificiale se manifestă și în arte, inclusiv în cinematografie. Se sărbătoresc 20 de ani de la premiera filmului *Avatar*, pelicula care în 2009 părea dintr-o altă lume grație efectelor speciale hiperrealiste, generate pe calculator. Pentru a sărbători această aniversare, Cameron dezvăluie un proiect secret bazat pe inteligență artificială: un al șaselea film *Avatar* bazat pe o tehnologie de detectare a mișcării subacvatice pe care o dezvoltase anterior, precum și un nou mediu virtual și un sistem de proiecție pe retină. Experiența este construită folosind algoritmi de generare meniți să creeze lumi complet noi pe care să le exploreze avatarurile umane, algoritmi evolutivi pentru randări și învățare profundă pentru efectuarea tuturor calculelor necesare. Rezultă un film unic prezentat într-un context special care (împreună cu sistemul de proiectare pe retină) oferă o experiență narativă complet originală-și care introduce spectatorul complet în subiect.

Inteligența artificială ajută organizațiile de orice dimensiune să aibă abordări de management mai creative. Mafia G oferă modele predictive pentru analiza de afaceri, oferind modalități de eficientizare, reducerea costurilor și domenii care ar putea fi îmbunătățite. Departamentele de resurse umane folosesc recunoașterea tiparelor pentru a evalua productivitatea și moralul angajaților- și pentru a elimina prejudecățile la angajare sau promovare. Nu mai există CV-uri; DDP-urile noastre ne arată punctele forte și slăbiciunile, iar programele de inteligență artificială ne analizează datele înainte de a ne recomanda unor posibili angajatori.

În cadrul multor companii mari, muncitorii umani nu mai prestează activități cu nivel cognitiv scăzut, în timp ce inteligența artificială sprijină angajații în anumite domenii legate de cunoaștere. Activitatea recepționerilor, a asistenților de clienți, a agenților de turism este acum pe deplin automatizată. În cadrul întâlnirilor, difuzoarele inteligente ascultă, folosind algoritmii de detectare a vocii pentru a transcrie conversațiile noastre. O secretară virtuală ia notițe automat, subliniind numele vorbitorilor, punctele importante discutate, zonele de convergență și divergență, informații contextuale din alte întâlniri și alte date relevante ale companiei. Sistemul stabilește temele care trebuie monitorizate ulterior și creează liste de activități pentru participanți.

Pentru că am recunoscut încă de la bun început că automatizarea va bulversa anumite părți ale forței de muncă, șomajul nu este răspândit iar economia noastră are fundații puternice. În Statele Unite, guvernul federal pune la dispoziție programe de recon-

versie profesională pentru a se asigura că suntem bine pregătiți. Folosind instrumentele Mafiei G, companiile și persoanele s-au pregătit din timp pentru noile tipuri de slujbe.

Mafia G a pus la dispoziția școlilor publice și private inteligența artificială pentru a spori procesul de învățare. Sistemele de învățare capabile de adaptare, supravegheate de profesori, îi fac pe elevi să învețe în propriul ritm, în special în ce privește cititul, logica, matematica și limbile străine. În sălile de clasă și în case, IBM l-a reînviat pe filosoful Socrate sub forma unui agent de inteligență artificială cu care dialogăm și care ne pune întrebări și ne răspunde pentru a ne stimula gândirea critică. Sistemul de inteligență artificială Socrate, evoluat din Watson, pune întrebări studenților legate de materie, dezbătând și discutând conceptele. (Inteligența artificială Socrate este folosită și în afara școlii, fiind un membru apreciat al oricărei echipe medicale, juridice, de ordine publică, strategie și politică. Ea este folosită și pentru a pregăti candidații pentru dezbaterile publice.)

Inteligența artificială Socrate a IBM e un aliat util în redacțiile de știri, ajutând jurnaliștii să analizeze știrile și posibile perspective alternative. Ea este folosită și pentru verificarea datelor și asigurarea calității editoriale: articolele sunt verificate pentru a nu conține prejudecăți inconștiente și a se asigura că se folosește un amestec variat de surse și perspective. (Au apus de mult timp vremurile în care ziarele și revistele publicau liste doar cu bărbați gânditori eminenți, oameni de afaceri de succes și așa mai departe.) Algoritmii de generare sunt folosiți pentru a crea videoclipuri din cadre foto-

grafice, a modela tridimensional peisaje şi clădiri pe baza câtorva fotografii şi a detecta voci distincte ascunse în mulţimi. Prin urmare apar mult mai multe reportaje video care sunt mai ieftin de produs.

Inteligenţa artificială este folosită pentru a detecta tipare şi anomalii ale datelor, făcând jurnaliştii să scoată la iveală noi relatări în interesul public. În loc să fie un complice al roboţilor de dezinformare, inteligenţa artificială poate detecta propaganda, afirmaţiile false şi campaniile de dezinformare. Ca urmare, democraţiile sunt mai puternice.

Mafia G a studiat oraşele din China în care se desfăşoară proiecte pilot ale oraşelor inteligente -cum ar fi Rongcheng, Beijing, Shenzhen, Shanghai- şi a identificat cele mai bune practici care ar putea fi preluate şi în Statele Unite. Avem acum câteva oraşe inteligente în America-Baltimore, Detroit, Boulder şi Indianapolis-care testează o gamă largă de servicii şi sisteme de inteligenţă artificială. Reţelele de sateliţi cubici-mici sateliţi de mărimea unui cub Rubik- transferă date în timp real sistemelor de inteligenţă artificială care pot recunoaşte obiecte, tipare unice de lumină şi amprente termice. Aceasta le permite edililor să prevadă eventuale pene de curent, să monitorizeze şi redirecţioneze traficul, să gestioneze rezervele de apă şi să deszăpezească drumurile. Inteligenţa artificială îi ajută şi să gestioneze bugetele şi personalul pe parcursul anului, descoperind modalităţi complet noi de a economisi la scară largă. Cheltuielile neprevăzute nu dispar, dar impactul lor se diminuează-iar cetăţenii acestor oraşe sunt animaţi de o senzaţie de speranţă pe care nu au mai trăit-o de mulţi ani.

Aceste sisteme furnizează date și autorităților de ordine publică, precum poliția sau pompierii, care folosesc inteligența artificială pentru a analiza cantități uriașe de date, inclusiv fluxuri video: dacă nu există sunet, algoritmii de recunoaștere a tiparelor pot citi de pe buze și pot transcrie conversațiile. Algoritmii de generare completează și ei lipsurile datelor audio iar dacă o înregistrare este neclară, un algoritm o poate face mai clară. Inteligența artificială scanează milioane de imagini, căutând tipare pe care un ochi omenesc le-ar putea rata. Instalarea acestor sisteme nu s-a făcut fără controverse. Cu toate acestea, angajamentele Mafiei G în domeniul protecției intimității fac ca DDP-urile noastre să nu poată fi accesate fără mandat. Avem siguranța că Mafia G ne apără intimitatea.

Pe măsură ce evoluează, inteligența artificială ne ajută să devenim niște oameni mai buni. Prin rolul activ jucat de Mafia G, guvernul federal și GAIA, ne simțim confortabil cu noțiunea de sugestionare.

2049: Rolling Stones au murit (dar fac muzică nouă)

Până în anii 2030, cercetătorii care lucrează în cadrul Mafiei G au publicat un articol stimulant, atât prin prisma a ceea ne dezvăluie despre inteligența artificială, dar și a felului în care a fost scris. Pornind de la același set de standarde și având sprijinul unei finanțări (și răbdări) substanțiale a guvernului federal, cercetătorii au colaborat pentru evoluarea inteligenței artificiale. Ca urmare, a fost dezvoltat primul sistem capabil de inteligență artificială generală.

Sistemul a trecut testul Colegului. A durat mult timp ca în comunitate să se accepte faptul că testul Turing, și alte teste similare, nu reprezintă barometrul corect pentru măsura inteligența mașinilor. Testele care se bazează fie pe amăgire (poate un computer să păcălească un om că discută că un om?) sau pe imitare (poate un computer să se comporte la fel cum am face-o noi?) nu recunosc particularitatea inteligenței artificiale: o inteligență obținută și exprimată într-o manieră care nu se aseamănă cu experiența umană. În loc să judecăm o IAG în funcție de capacitatea acesteia de a „gândi" la fel cum o facem noi, comunitatea inteligenței artificiale a adoptat în sfârșit un nou test pentru a măsura *contribuțiile semnificative* ale unei IAG, care să judece valoarea sarcinilor cognitive și comportamentale- diferite, dar puternice- pe care nu le-am putea efectua pe cont propriu. IAG se consideră atinsă atunci când sistemul poate contribui în mod general de o manieră similară sau superioară unui om.

Mafia G a cercetat și dezvoltat timp de ani de zile o IAG care să poată asista la o ședință de lucru și să contribuie cu o sugestie valoroasă-fără ca să i se ceară acest lucru- înainte de finalul ședinței. Această IAG a primit numele de cod „Proiectul Hermione", după personajul din *Harry Potter* care știa întotdeauna și în orice situație ce să spună sau să facă. Majoritatea oamenilor de pe glob au adus contribuții valoroase în cadrul unui grup, la un moment dat: la muncă, într-un context religios, la barul din cartier cu prietenii sau într-un curs de istorie din liceu. Simpla întrerupere doar de dragul de a prezenta o informație sau de a răspunde la o întrebare nu sporește valoarea unei conversații. O contribuție însemnată presupune o combinație a mai multor aptitudini:

Capacitatea de a face presupuneri informate: Ea se mai numeşte şi „raţionament pe furiş" şi reprezintă felul prin care reuşim să ne descurcăm în viaţă. Folosim cele mai bune informaţii pe care le avem, enunţăm şi verificăm ipoteze şi găsim răspunsuri chiar şi în absenţa unor explicaţii clare.

Extragerea corectă a semnificaţiei cuvintelor, pauzelor şi zgomotului de fond: Doar pentru că cineva spune că e „încântat" să se ocupe de un nou proiect nu înseamnă că e *încântat* la propriu. Alte indicii, precum limbajul lor corporal, ne-ar putea transmite că sunt destul de *neîncântaţi* să accepte cererea, însă nu sunt în stare să spună nu.

Folosirea experienţei, cunoaşterii şi a contextului istoric pentru a înţelege: Atunci când oamenii interacţionează, ei aduc cu sine o perspectivă nuanţată asupra lumii, un set unic de experienţe personale şi, de regulă, şi propriile aşteptări. Uneori, logica şi datele nu sunt suficiente pentru a câştiga o dezbatere. Alteori, ele sunt singurele aspecte care contează.

Interpretarea încăperii: Există o interacţiune explicită-şi una tacită, care se petrece în profunzime. Indiciile subtile ne pot ajuta să ne dăm seama atunci când este o problemă conside rabilă care trebuie să ne capteze atenţia.

Proiectul Hermione a participat la o sesiune a unui grup de lucru al GAIA. Optsprezece membri ai grupului au discutat şi dezbătut standardele existente pentru inteligenţa artificială, dezvoltate fie de către oamenii aflaţi în încăpere, fie de predecesorii acestora. Deoarece grupul era divers şi alcătuit din lideri din diferite ţări şi culturi, există multe informaţii subliminale: anumite dinamici de putere, conflicte de personalitate şi senzaţii de inferioritate sau su-

perioritate. Grupul a discutat cu IAG ca un membru egal, fără beneficii suplimentare sau excepții speciale. Pe la jumătatea sesiunii, IAG a respins un consens ușor majoritar, dar cu potențial de creștere, care se degajase în jurul temei reglementărilor. IAG a argumentat cu tact împotriva ideii și a obținut sprijinul unui alt membru al grupului pentru a susține o alternativă. Proiectul Hermione a făcut o contribuție importantă. (De neprețuit, ar putea susține alții ulterior.)

Proiectul Hermione a fost un succes nu doar pentru că a reușit să treacă testul Colegului cu atâta ușurință-ci și din cauză că GAIA și Mafia G au văzut în acest moment atât un semnal de alarmă cât și o oportunitate. Aceste organizații au continuat să-și recalibreze strategiile și standardele pentru a putea ține pasul cu evoluția tehnologică a inteligenței artificiale. Organizațiile au decis să limiteze ritmul auto-îmbunătățirii, adăugând constrângeri tuturor sistemelor bazate pe inteligență artificială pentru ca oamenii să fie tot timpul informați. Cercetătorii GAIA urmează acum noi protocoale: ei rulează simulări pentru a putea înțelege impactul unor IAG mai puternice înainte de a le permite să fie folosite în scopuri generale, comerciale sau militare.

Membrii Mafiei G sunt companii bogate, puternice și cu influență-iar succesul lor sporește. Aceste companii construiesc aplicații practice interesante ale IAG pentru a ne spori productivitatea și creativitatea, contribuind și la oferirea de soluții pentru problema cea mai presantă a omenirii: schimbările climatice. Pe măsură ce curentul Atlantic s-a deplasat înspre nord, și grânarul Americii a urmat același drum, înspre Canada, decimând fermele și agricultura Statelor Unite. Cafeaua și ciocolata nu mai pot fi

crescute cu ușurință în aer liber. Cetățenii din Bangladesh, Filipine, Tailanda și Indonezia au devenit refugiați climatici în propriile țări. Amazon, în parteneriat cu Microsoft, grupul francez Danone și DowDuPont din Statele Unite folosesc IAG și experimentele genetice pentru a crește legume în sere.

Google și Facebook folosesc IAG pentru a muta întregi populații în siguranță, creând noi comunități omenești. IAG îi ajută să găsească locurile în care pot trăi cel mai ușor o viață confortabilă și care să păstreze cultura popoarelor afectate. Regiunile altădată pustii ale planetei sunt modelate sau transformate folosind materiale de construcții adaptabile. Zgârie-pământurile- complexe mari, întinse, de doar câteva etaje înălțime- modifică complet amprenta urbană. În interior, lifturi fără cabluri ne transportă în orice direcție. E o tendință nouă în arhitectură care ajută cele mai importante centre economice ale lumii să prospere -inclusiv orașele americane Denver, Minneapolis și Nashville.

O perioadă părea că China se va retrage și se va regrupa în jurul puținilor aliați rămași- Coreea de Nord, Rusia, Mongolia, Myanmar, Cambogia, Kazahstan, Pakistan, Kîrgîzîstan, Țadjikistan și Uzbekistan. Universitățile din țările GAIA nu mai acceptă studenți din China. Din frica de supraveghere și pierdere a DDP-urilor, industria turismului dispare din China. Țările GAIA se bazează pe sisteme automate care să producă materiile prime pentru fabricație, reluând fabricarea produselor în țară. În cele din urmă, guvernul chinez a stabilit că excluderea din GAIA era nocivă pentru economie, provocând tulburări politice și economice conside-

rabile. China a acceptat cu o oarecare reticență să adopte normele și standardele GAIA, precum și măsurile de transparență cerute de statele membre GAIA. Comunismul nu a murit- există încă destule puncte politice fierbinți, dar și tensiunile obișnuite legate de diferitele stiluri de guvernare și conducere.

IAG nu a apărut fără multe probleme noi, unele putând fi anticipate. Ca și alte tehnologii care au transformat cu timpul societatea omenească, IAG a eliminat slujbe, a condus la apariția unor noi tipuri de activități ilicite și a scos la iveală partea urâtă a noastră. În anii 2040, IAG nu este însă o amenințare existențială.

Folosim IAG atât la muncă cât și acasă pentru a accesa informații. IAG este un agent de control care ia multiple forme, în funcție de situație: vorbim cu ea, interacționăm cu ea pe un ecran și îi trimitem date provenite din corpurile noastre. Toate familiile au un valet, pentru că în fiecare casă există o IAG antrenată și calibrată în funcție de împrejurările casei.

Una dintre cele mai mari și mai notabile schimbări aduse de IAG este o sporire drastică a sofisticării în majoritatea aspectelor existenței umane. Putem să fim recunoscători Mafiei G pentru creșterea considerabilă a calității vieții. Activități dificile, cronofage -cum ar fi găsirea unui timp convenabil pentru toată lumea, stabilirea unui program de activități după școală, gestionarea finanțelor personale- sunt acum complet automatizate și supravegheate de către IAG. Nu ne mai pierdem ore întregi încercând să facem curat în căsuța poștală- IAG ne înlesnește sarcinile cognitive facile. Avem în sfârșit aparate electrocasnice care ne curăță pode-

lele şi covoarele, ne pun rufele la uscat şi ne şterg praful. (Ne aducem aminte de 2019 ca de o epocă mult mai simplă, plină de sarcini manuale monotone, plictisitoare.)

Răceala nu mai există, şi nici gripa. De fapt, ne mirăm acum cât de naivi erau doctorii din trecut. ŞI asta pentru că IAG ale IBM şi Google ne-au ajutat să vedem şi înţelegem milioane de viruşi diferiţi. Acum, când nu ne simţim bine, o IAG efectuează un test de diagnostic pentru a stabili cauza astfel încât să poată fi prescris un tratament în funcţie de DDP. Au dispărut şi medicamentele fără reţetă, însă preparatele magistrale cunosc o revenire. Aceasta se datorează faptului că IAG a contribuit la accelerarea unor evoluţii critice în domeniul modificărilor genetice şi al medicinei de precizie. Puteţi apela acum la serviciile unui farmacist informatic: farmaciştii special pregătiţi cu pregătire în bioinformatică, medicină şi farmacologie.

Farmacia informatizată este o specializare medicală separată, care colaborează îndeaproape cu un nou tip de medici de familie cu formare în inteligenţă artificială: medici generalişti cu studii de medicină şi tehnologie. Deşi IAG a eliminat anumiţi specialişti: radiologii, imunologii, alergologii, cardiologii, dermatologii, endocrinologii, anesteziştii, neurologii şi alţii- doctorii care activau în aceste domenii au avut suficient timp la dispoziţie pentru a dobândi noi competenţe în domenii adiacente. Viaţa pacientului este mai bună. Nu mai petreceţi ore în şir de la un doctor la altul, primind mesaje contradictorii şi nu vi se mai prescriu medicamente în exces. Dacă trăiţi într-o zonă mai izolată, apariţia IAG v-a îmbunătăţit semnificativ accesul la îngrijirea medicală.

De la naştere, fiecărui copil i se scanează genomul- acest procedeu este rapid şi la îndemâna oricui. Oamenii preferă să aleagă această variantă pentru că secvenţa genetică reprezintă o componentă esenţială a DDP-ului. Pe lângă faptul că vă oferă o fereastră prin care puteţi privi alcătuirea voastră genetică unică, IAG analizează datele pentru a detecta anomalii genetice şi a afla mai multe despre cum funcţionează corpul vostru. Desigur, atât în Statele Unite cât şi în alte ţări există în continuare minorităţi care se opun acestei practici- aşa cum existau cândva oameni care se opuneau vaccinării. Deşi părinţii pot alege să se opună din motive religioase sau ideologice, puţini aleg această cale.

Graţie IAG suntem mai sănătoşi-şi avem la dispoziţie noi opţiuni când vine vorba de căutarea unui partener şi de căsătorie. Forme avansate de intimitate diferenţiată permit unui terţ să acceseze datele (DDP-ul, genomul şi istoricul medical) fără ca să ştie cui aparţin aceste date. Furnizorii de servicii matrimoniale care au implementat tehnologia IAG devin astfel deosebit de utili, pentru că acum putem alege să optimizăm în funcţie de viitoarea familie (copii din combinaţii genetice dezirabile), avere (estimarea potenţialului de câştig de-a lungul vieţii) sau distracţie (dacă vor râde sau nu la glumele voastre).

IAG vă ajută şi în alte demersuri creative, dincolo de căutarea iubirii. Membrii Rolling Stones au murit cu mulţi ani în urmă, dar

grației algoritmilor mimetici, ei continuă să creeze muzică. Aceeași senzație pe care ați avut-o când ați auzit prima dată primele 30 de secunde ale melodiei „Paint It Black" -melodia melancolică de chitară, urmată de 8 bătăi puternice de tobe și un refren repetitiv în care Mick Jagger cântă „I see a red door and I want it painted black..." – un moment unic de stimulare și satisfacție. Nu părea posibil că veți putea simți același lucru din nou față de un cântec al Rolling Stones, și cu toate acestea cea mai nouă melodie a lor este la fel de zgomotoasă, de dură și de stimulantă.

Deși ziarele pe hârtie au dispărut, presa a adoptat IAG ca metodă de distribuție. Odată ce IAG au putut trece testul Colegului, redacțiile au reacționat rapid pentru a construi un nou model de distribuție a știrilor, care să poată fin înc continuare profitabil dar să țină mai mult cont de preferințele viitorului. În zilele noastre, majoritatea oamenilor nu „primesc" știrile sau nu „dau drumul" la știri- ei au o discuție cu un agent de știri inteligent. *New York Times* și *Wall Street Journal* angajează sute de jurnaliști informatici-oameni cu pregătiri hibride în inteligența artificială și jurnalism. Împreună, aceste echipe relatează știri și aleg datele relevante care trebuie incluse în motoarele de conversație. Jurnalismul bazat pe IAG ne informează, și putem să-l configurăm pentru a-i da o culoare politică, a primi mai multe date secundare sau o variantă extinsă cu personaje secundare și fapte diverse. Suntem consultați în ce privește analiza știrilor și a fluxurilor editoriale, dezbătând în mod constructiv cu agentul de știri folosind vocea sau interacționând cu ecranele (ochelari inteligenți și tablete retracta-

bile). Există în continuare articole de amploare, atât sub formă text cât și video.

Hackerii IAG-în mare parte *alte IAG*- sunt o problemă constantă pentru că aceste IAG comit ilegalități, dezvăluind oamenii care le-au creat codul sursă. Agențiile guvernamentale angajează ofițeri cu experiență în analiza datelor. Cu ajutorul grupului BAT din China, cei Nouă Titani colaborează la dezvoltarea unei aparaturi, a unor structuri, rețele și algoritmi capabile să reziste atacurile. Parteneriatul dintre GAIA și INTERPOL a reușit în mare măsură să limiteze criminalitatea organizată.

Proiectele pilot de orașe inteligente lansate în urmă cu două decenii în Baltimore, Indianapolis, Detroit și Boulder au fost de succes și au ajutat și alte comunități să-și dezvolte bunele practici, conducând la înființarea Administrației Federale a Infrastructurii Inteligente. Ca și Administrația Federală a Autostrăzilor, AFII operează sub egida Departamentului Transporturilor și supraveghează toate sistemele conectate care ne alimentează orașele: stații de transformare fără cabluri, generatoare descentralizate de energie (cinetice, solare și eoliene), rețele de conectare a vehiculelor la infrastructură, fibra optică prin care se aduce lumina soarelui în fermele subterane. Datele senzoriale sunt adunate și folosite pentru a modela sănătatea generală a comunităților noastre: accesul la un aer curat, curățenia cartierelor, folosirea parcurilor și a zonelor exterioare de agrement. IAG prevede și elimină penele de curent și penuriile de apă înainte de a se produce.

Pe măsură ce ne apropiem de tranziția de la IAG la SIA, se întrevede o oportunitate extraordinară: interfețele între creier și mașină. Ne aflăm pe pragul de a deține nanotehnologia moleculară, sperând ca în câteva decenii să putem înregistra simultan datele generate de miliardele de neuroni din creierul nostru. Computere microscopice, de dimensiunea unui fir de nisip, ar urma să se stea în vârful creierului și să capteze semnalele electrice. Sisteme IAG speciale, capabile să citească și să interpreteze aceste semnale, ar putea de asemenea să transmită date între oameni. O interfață creier-mașină ar putea cândva permite unei persoane sănătoase să redea funcționalitatea unei persoane care a suferit un accident vascular cerebral, o paralizie sau care nu mai poate vorbi. Interfețele creier-mașină, pe care le-am putea folosi teoretic pentru a transfera amintiri între oameni, ne-ar putea ajuta să trăim empatia de o manieră mai profundă și mai plină de semnificație.

Acest potențial ne face să ne gândim la noi utilizări ale IAG. Vrem să aflăm răspunsul la întrebări filosofice spinoase: *Trăim într-un univers real? E posibil ca „nimicul" să existe? Care e natura timpului?* IAG nu ne poate da răspunsurile pe care ni le dorim, însă Mafia G ne poate spori înțelegerea condiției umane.

2069: Gardienii Galaxiei cu puterea inteligenței artificiale

Explozia inteligenței, anticipată în urmă cu 100 de ani de matematicianul și pionierul inteligenței artificiale I.J. Good, începe la finalul anilor 2060. Devine clar că IAG atinge niveluri profunde de inteligență, viteză și putere, și că super-inteligența artificială

poate fi atinsă în viitorul apropiat. În ultimul deceniu, cei Nouă Titani și GAIA s-au pregătit pentru acest eveniment- calculând că odată ce nivelul inteligenței omenești va fi depășit de mașini, SIA ar putea să apară după doar câțiva ani.

După multe dezbateri, membrii GAIA iau decizia dificilă de a împiedica apariția unei SIA. Unii dintre cei care au participat s-au emoționat, argumentând că nu e just ca „mințile strălucite" ale inteligenței artificiale să fie mutilate taman când se apropiau de potențial. Dezbatem dacă negăm sau nu omenirii posibilitatea de a avea câștiguri și oportunități chiar mai mari.

În cele din urmă, cu binecuvântarea și încurajarea celor Nouă Titani, GAIA stabilește că toate IAG trebuie să aibă restricții care să le limiteze rata auto-îmbunătățirii și care să împiedice apariția unor mutații nedorite, pentru a apăra siguranța și securitatea oamenilor. În curând, GAIA va lansa o serie de paznici bazați pe inteligență artificială, care să acționeze ca un sistem de avertizare timpurie pentru orice IAG care capătă prea multă putere cognitivă. Chiar și gardienii nu vor împiedica o entitate ilicită să încerce să creeze propriile SIA, însă GAIA scrie scenarii pentru a putea contracara o astfel de situație. GAIA și cei Nouă Titani beneficiază de afecțiunea și încrederea noastră nelimitate.

CAPITOLUL ŞASE

CUM SĂ ÎNVĂȚĂM SĂ TRĂIM CU MILIOANE DE TĂIETURI CU HÂRTIA: SCENARIUL PRAGMATIC

Până în 2023, problemele inteligenței artificiale sunt cunoscute dar decidem pe parcurs să modificăm doar în mică măsură parcursul de dezvoltare al inteligenței artificiale, un sistem despre care ştim cu toții că prezintă defecte evidente. Urmărim doar mici retuşuri pentru că cei care controlează inteligența artificială nu sunt dispuşi să-şi asume măsuri incomode: să renunțe la potențialele beneficii financiare, să ia decizii politice nepopulare, şi să ne tempereze aşteptările exagerate pe termen scurt, chiar dacă asta ar însemna îmbunătățirea şanselor de a trăi pe termen lung alături de inteligența artificială. Chiar mai grav este faptul că ignorăm China şi planurile sale de viitor.

Liderii din Congres, din diversele agenții federale şi de la Casa Albă continuă să nu acorde o prioritate prea mare inteligenței artificiale şi cercetării ştiințifice avansate în general, preferând să investească în industrii cu mai mare priză din punct de vedere politic dar care sunt aproape de finalul vieții. Planul strategic pri-

vind viitorul inteligenței artificiale publicat de administrația Obama în 2016- un document care a influențat puternic planul strategic chinez pentru anul 2025- este abandonat, împreună cu programul federal de finanțare a cercetării și dezvoltării în domeniul inteligenței recomandat de acest plan. America nu are o viziune sau o strategie pe termen lung în domeniul inteligenței artificiale, și nu crede că aceasta va avea vreun impact economic, educațional sau asupra securității naționale. Politicienii americani, din ambele partide, se concentrează asupra înăbușirii Chinei, în loc să dezvolte strategii pentru a coaliza forțele Mafiei G și ale guvernului. Absența unei coaliții și a unei strategii naționale coerente în privința inteligenței artificiale generează tăieturi cu hârtia- milioane și milioane- care încep să sângereze pe parcurs. Ele nu se observă de la început. Cultura pop, articolele jurnaliștilor din domeniul tehnologiei și postările de pe rețelele de socializare ale persoanelor influente ne-au învățat să căutăm semne mari, evidente- cum ar fi roboți ucigași- prin urmare, vom rata adevăratele semnale de alarmă, mai mici și mai puțin evidente, pe măsură ce inteligența artificială evoluează. Cei Nouă Titani sunt forțați să acorde prioritate vitezei în detrimentul siguranței, astfel încât parcursul de dezvoltare al inteligenței artificiale – de la IAÎ către IAG și mai departe- continuă fără să rezolve mai întâi vulnerabilitățile tehnice grave. Iată câteva dintre tăieturile cu hârtia mai puțin evidente- multe fiind auto-induse- pe care nu le tratăm ca niște răni serioase, așa cum vor ajunge să devină.

În calitate de consumatori de tehnologie, avem așteptarea ca triburile inteligenței artificiale să-și fi imaginat și să fi rezolvat toate

problemele înainte ca aplicațiile, produsele sau serviciile să părăsească laboratoarele de cercetare și dezvoltare. Ne-am obișnuit să adoptăm o tehnologie care funcționează imediat ce o scoatem din cutie. Când cumpărăm noi telefoane inteligente și televizoare, le băgăm în priză, iar ele funcționează așa cum ni se promite. Când descărcăm programe noi, fie că sunt procesoare de text sau programe de analiză a datelor, ele se comportă conform așteptărilor. Uităm că inteligența artificială nu este o tehnologie care să funcționeze imediat ce o scoatem din cutie, pentru că pentru a putea funcționa așa cum ne dorim, un sistem de inteligență artificială are nevoie de cantități uriașe de date și de ocazia de învăța în timp real.

Niciunul dintre noi-consumatori, jurnaliști sau analiști- nu permite celor Nouă Titani să facă greșeli. Pretindem noi produse, servicii, invenții sau descoperiri, și ne lamentăm în public când nu le primim. Nu ne interesează că solicitările noastre distrag atenția triburilor inteligenței artificiale de la perfecționarea produselor.

Modelele și structurile inteligenței artificiale, indiferent cât de mari sau de mici, au nevoie de multe date pentru a putea învăța, a se îmbunătăți și a putea fi implementate. Datele sunt ca oceanele lumii. Ne înconjoară, constituie o resursă nelimitată, și complet inutilă dacă nu o desalinizăm, tratăm și procesăm pentru a o putea consuma. În momentul de față există doar câteva companii care pot desaliniza eficient datele la o scară care să conteze. Acesta este motivul pentru care partea cea mai dificilă a construirii unui nou sistem de inteligență artificială nu se rezumă la algoritmi sau la modele, ci la colectarea datelor potrivite și etichetarea lor corectă

pentru ca maşina să se poată antrena şi învăţa de pe urma lor. Comparativ cu diversitatea produselor şi serviciilor pe care cei Nouă Titani se străduie în grabă să le construiască, există foarte puţine seturi de date care sunt gata de a fi folosite. Câteva dintre acestea sunt ImageNet (un set uriaş de imagini folosit la scară largă), WikiText (un set de date pentru modelarea limbajului bazat pe articole din Wikipedia), 2000 HUB5 (un set de date disponibil doar în engleză folosit pentru vorbire) şi LibriSpeech (aproximativ 500 de ore de mostre din audiobook-uri. Dacă aţi dori să construiţi o inteligenţă artificială pentru sănătate care să detecteze anomalii în analizele de sânge sau scanările oncologice, problema nu e inteligenţa artificială, ci lipsa datelor- oamenii au organisme complicate, cu multe variaţii, şi nu există un set suficient de mare de date gata de a fi folosite.

În urmă cu un deceniu, în anii 2010, echipa IBM Watson Health a încheiat un parteneriat cu mai multe spitale pentru a vedea dacă inteligenţa artificială ar putea sprijini activitatea doctorilor. Watson Health avea la activ nişte realizări remarcabile, inclusiv un caz în care era implicat un băieţel bolnav în vârstă de 9 ani. După ce specialiştii nu au putut să-l diagnosticheze şi să-l trateze, Watson a atribuit o probabilitate mai multor probleme de sănătate- pe listă fiind atât boli mai frecvente cât şi mai exotice, inclusiv un sindrom rar întâlnit în copilări denumit boala Kawasaki. Odată ce s-a auzit că Watson pune diagnostice miraculoase şi salvează vieţi omeneşti, echipa Watson a fost supusă unor presiuni de a comercializa platforma, fiind stabilite ţinte nerealiste. IBM proiectase că Watson

Health avea să crească de la venituri de 244 de milioane de dolari în 2015, la 5 miliarde de dolari până în 2020.[175] Aceasta ar fi reprezentat o creștere de 1949% în mai puțin de cinci ani.

Înainte ca Watson Health să poată să stârnească din nou aceeași senzație ca în trecut-respectând în același timp și un program extrem de ambițios de dezvoltare- sistemul ar fi avut nevoie mult mai multe date și de timp pentru a învăța. Nu existau însă suficiente date de sănătate din lumea reală, iar cele existente nu erau suficient de amănunțite. Motivul era că aceste date erau stocate în arhive gestionate de o altă companie, care considera că IBM îi face concurență.

Ca urmare, echipa IBM a folosit o variantă ocolitoare frecventă întâlnită în rândul triburilor inteligenței artificiale. Sistemul Watson Health a fost alimentat cu așa-zise „date sintetice", respectiv date care reprezintă informații ipotetice. Deoarece cercetătorii nu dispun de date brute pe care să le introducă într-un sistem de învățare pentru a-l pregăti, ei cumpără un set de date sintetice de la o terță firmă sau construiesc chiar ei unul. Această abordare este adesea problematică pentru că alcătuirea acelui set de date- ce date trebuie incluse, cum trebuie etichetate- depinde de deciziile unui număr redus de oameni, care adesea nu sunt conștienți de propriile prejudecăți politice, profesionale, de gen, și altele.

Așteptările exagerate vizavi de profitabilitatea imediată a Watson Health, combinate cu o dependență de seturile de date sintetice au condus la o problemă gravă. IBM încheiase un parteneriat cu Memorial Sloan Kettering Cancer Center pentru a folosi Watson Health în domeniul oncologic. La scurt timp după aceea, câțiva experți care colaborau la proiect au relatat exemple de recomandări incorecte sau hazardate de tratament. Într-un astfel de

caz, Watson Health a recomandat un protocol bizar de tratament unui pacient diagnosticat cu cancer pulmonar care manifesta și semne de hemoragie: chimioterapie și un medicament denumit bevacizumab, care ar fi fost contraindicat pentru că poate provoca hemoragii grave sau chiar fatale. [176] Vestea că Watson e inept s-a răspândit în publicațiile de specialitate din domeniul sănătății și pe blogurile informaticienilor, fiind adesea acompaniată de titluri incendiare. Cu toate acestea, problema nu consta în faptul că Watson Health voia să le facă oamenilor de hac, ci că forțele de piață au exercitat presiuni asupra IBM pentru a grăbi ritmul cercetării în inteligența artificială pentru a îndeplini planurile financiare.

Iată o altă tăietură cu hârtia: unele inteligențe artificiale și-au dat seama cum să trișeze și să-și modifice propriile sisteme. Dacă o inteligență artificială este programată cu scopul de a învăța un joc, a-l juca și a face orice e nevoie pentru a câștiga, cercetătorii au descoperit cazuri în care recompensele au fost manipulate, prin faptul că un sistem se va folosi de algoritmi evolutivi și de învățare pentru a câștiga folosind înșelătoria și disimularea. Spre exemplu, o inteligență artificială care joacă Tetris și-a dat seama că ar putea pur și simplu să pună jocul pe pauză pe termen nelimitat pentru a nu pierde niciodată. Încă de când ați citit prima dată despre manipularea recompenselor -știrea a devenit celebră recent când două sisteme de inteligență artificială din domeniul financiar au prezis o scădere dramatică a bursei și au încercat să închidă pe cont propriu bursele pe termen nelimitat- v-ați întrebat ce s-ar putea întâmpla dacă datele voastre ar fi exploatate într-un sistem

de manipulare a recompenselor. Concediul acele din vacanța de iarnă care urmează- cum ar fi dacă serviciile de control al traficului aerian nu ar mai funcționa?

O altă tăietură cu hârtia: agenții răuvoitori pot introduce date otrăvite în programele de pregătire a inteligenței artificiale. Rețelele neurale sunt vulnerabile la „exemplele proaste" care sunt informații fictive sau eronate menite să inducă intenționat un sistem de inteligență artificială să facă o greșeală. Un sistem de inteligență artificială ar putea eticheta o imagine ca fiind un panda, cu un grad de siguranță de 60%, dar dacă adăugăm un pic de distorsiuni imaginii, câțiva pixeli pe care un om nu i-ar detecta, iar sistemul ar putea re-eticheta imaginea ca un gibon, cu o siguranță de 99%. E posibil să păcălim sistemul de vizualizare digitală al unei mașini pentru a crede că un semafor roșu transmite de fapt „limită de viteză 100", iar pasagerii să treacă la viteză maximă printr-o intersecție. Datele eronate ar putea reprograma un sistem militar de inteligență artificială să interpreteze semnele unui spital obișnuit -ambulanțe, cuvântul „urgență" sau „spital" ca aparținând unor organizații teroriste. Problema e că cei Nouă Titani nu și-au dat seama cum ar putea să-și protejeze sistemele de exemplele proaste, nici în lumea digitală, nici în cea fizică.

O tăietură mai profundă: cei Nouă Titani știu că informațiile false pot fi folosite pentru a reprograma sistemele automate de

învățare și rețelele neurale. O echipă din cadrul diviziei Brain a Google a publicat în 2018 un articol prin care explica cum un agent malefic ar putea introduce informații false într-o bază de date dedicată vizualizării digitale și să reprogrameze practic tot ce învață sistemele din acea bază de date. [177] Hackerii ar putea într-o bună zi să introducă date otrăvitoare în căștile voastre inteligente și să le reprogrameze cu identitatea altei persoane doar prin redarea unor sunete nocive în preajma voastră în metrou.

Situația se complică prin faptul că informațiile contrare pot avea și o utilitate. O altă echipă a Google Brain a descoperit că informațiile contrare pot fi folosite într-o „rețea antagonică generativă" sau RAG, pe scurt. E ca un test Turing, dar fără niciun om. Două inteligențe artificiale sunt programate folosind aceleași date- niște imagini cu oameni, de pildă. Prima generează fotografii ale dictatorului nord coreean Kim Jong-Un care să pară realiste, în timp ce a doua compară fotografiile generate cu cele reale. Pe baza analizei celei de-a doua inteligențe artificiale, prima revine și rafinează propriul proces. Secvența se repetă până când prima inteligența artificială generează automat tot soiul de imagini cu Kim Jong-Un care par complet realiste, dar care nu au existat niciodată în lumea reală. Imagini care-l arată pe Kim Jong-un luând cina cu Vladimir Putin, jucând golf cu Bernie Sanders sau sorbind cocktailuri cu Kendrick Lamar. Google Brain nu urmărește subterfugiile, ci să rezolve problema datelor sintetice. RAG ar permite sistemelor de inteligență artificială să lucreze cu date brute, din lumea reală, care nu au fost filtrate și fără supravegherea directă a unui programator uman. Și deși e o abordare deosebit de creativă pentru rezolvarea unei probleme, ea ar putea într-o zi să constituie o amenințare gravă la adresa siguranței noastre.

Încă o tăietură cu hârtia: când algoritmii complecşi colaborează, ei concurează uneori unul împotriva celuilalt pentru a îndeplini un obiectiv, iar acest fapt poate otrăvi un sistem întreg. Am fost martorii unor probleme sistemice atunci când preţul unui manual de biologie a început să crească galopant. Cartea nu se mai tipărea, dar pe pagina Amazon se afişa că există 15 copii folosite de la alţi comercianţi, începând de la un preţ de 35,54 dolari- şi două copii noi noute la un preţ de 1,7 milioane de dolari. În spatele siteului, algoritmii Amazon au demarat un război al preţurilor, alegând să ridice preţul din ce în ce mai mult până când a ajuns la 23.698.655 dolari (plus 3 dolari şi 99 de cenţi pentru expediere). Sistemul algoritmilor de învăţare ajustase în timp real ca reacţie la fiecare licitaţie, iar aceste operaţii făceau parte din bagajul cu care fusese programat. Ca să reformulăm: e posibil să fi învăţat fără să vrem inteligenţa artificială că bulele de pe pieţe sunt un lucru bun. Nu e greu să ne imaginăm algoritmi rivali care umflă preţul caselor, acţiunilor sau chiar al publicităţii digitale până la cote absurde.

Acestea sunt doar o mică parte dintre tăieturile cu hârtia pe care triburile inteligenţei artificiale au hotărât că le putem suporta de dragul atingerii obiectivelor stabilite de forţele de pe piaţă în Statele Unite, respectiv de Partidul Comunist Chinez la Beijing. În loc să domolească pretenţiile legate de viteză şi profitabilitate, triburile inteligenţei artificiale sunt presate în continuu să scoată produse pe piaţă. Siguranţa e doar un aspect secundar. Angajaţii şi conducerea

Mafiei G sunt îngrijorați, însă nu le permitem suficient timp pentru a face schimbările necesare. ȘI încă nu am vorbit deloc de China.

Între 2019 și 2023 am ignorat efectiv dezideratele lui Xi Jinping pentru viitor: strategia națională amănunțită a Chinei în domeniul inteligenței artificiale, planurile lui de a domina economia globală, și obiectivul Chinei de a deveni o forță unică de decizie în domeniul geopolitic. Am eșuat în a lega viitorul inteligenței artificiale de infrastructura de supraveghere și sistemul de credit sociale, precum și activitatea diplomatică chineză în diferite țări africane, asiatice și europene. Așa că, atunci când Xi vorbește public și frecvent despre nevoia unei reforme a sistemului politic mondial și continuă înființând entități internaționale precum Banca Asiatică de Investiții în Infrastructură, nu îi acordăm toată atenția, ci doar o privire cu coada ochiului. Aceasta este o greșeală pe care nu o recunoaștem imediat.

În China, parcursul către o poziție dominantă în inteligența artificială nu a fost lipsit de obstacole. China are propriile tăieturi cu hârtia cu care se confruntă, deoarece grupul BAT se luptă să poată inova în același ritm cu Silicon Valley sub mâna forte a Beijingului. Grupul BAT eludează în mod repetat reglementările birocratice. Toate aceste scandarluri- cum ar fi cel prin care Administrația de Stat a Rezervelor Valutare din China a amendat Alipay cu suma de 600.000 de yuani (aproximativ 88.000 de dolari) pentru înregistrarea contabilă eronată a unor plăți internaționale între 2014 și 2016, în timp ce Tenpay a fost pedepsită pentru că nu a depus documentele corecte de înregistrare în vederea efectuării

de plăți internaționale între 2015 și 2017-s-au dovedit a nu reprezenta doar excepții.[178] Devine evident că acestea nu reprezintă incidente izolate iar oficialii chinezi trebuie să se confrunte cu tensiunile dintre abordările socialiste și realitățile capitalismului.

Vedem deja consecințele tuturor acestor vulnerabilități de ordin politic, strategic și tehnic. Pentru a face pe plac finanțatorilor, Mafia G preferă să încheie contracte profitabile cu guvernul în locul unor parteneriate strategice. Astfel, se stimulează un climat concurențial în locul colaborării. Consecința este o interoperabilitate redusă a structurilor, serviciilor și aparatelor care folosesc inteligența artificială. La începutul anilor 2020, piața a sugestionat Mafia G să-și împartă anumite funcționalități și activități: Amazon domină în casele noastre și în comerțul electronic, în timp ce Google domină în domeniul căutărilor pe internet, al localizării, comunicării personale și la locul de muncă. Microsoft domină în domeniul cloud, în timp ce IBM și-a arogat aplicațiile inteligenței artificiale la nivel de companii și sisteme aplicate de sănătate. Rețele de socializare sunt dominate de Facebook, iar Apple fabrică aparate (telefoane, computere, accesorii).

Niciun membru al Mafiei G nu aderă la un singur set de valori de bază care să acorde prioritate transparenței, incluziunii și siguranței. Deși conducătorii companiilor din Mafia G sunt de acord că ar trebui probabil ca în domeniul inteligenței artificiale să existe standarde adoptate la scară largă și ca ele să fie implementate, nu există pur și simplu o modalitate prin care să se aloce timp sau resurse pentru ele.

Dosarul de Date Personale este construit, întreținut și deținut de patru companii ale Mafiei G: Google, Amazon, Apple și Facebook. Iată însă și problema: nici măcar nu știți că aceste DDP-uri există sau că Mafia G și triburile inteligenței artificiale se folosesc de ele. Nu este o omisiune intenționată, ea datorându-se pur și simplu vitezei. Existența lor se regăsește în termenii și condițiile pe care le acceptăm întotdeauna, fără să le citim niciodată.

Formatul folosit de fiecare furnizor de DDP nu este complementar așa că există multe date duplicat care există în mai multe locuri, dar și, în mod paradoxal, multe spații goale unde lipsesc date. E ca și cum patru fotografi diferiți v-ar face o poză: una cu reflectoare și umbrele de reflecție, una cu un obiectiv ochi-de-pește, una cu o cameră de unică folosință și una cu un aparat RMN. Din punct de vedere tehnic, sunt patru imagini ale capului vostru, dar datele pe care le conțin cele patru sunt radical diferite.

Străduindu-se să obțină o imagine mai completă, triburile inteligenței artificiale trimit „ambasadori digitali" -mici programe care acționează ca intermediari și negociază în numele Mafiei G. Ambasadorii digitali de la Google și Amazon funcționează o perioadă, dar nu reprezintă soluții realiste pe termen lung. Aceștia sunt prea dificil de actualizat, mai ales că atât de multe produse și servicii oferite de terți sunt legate de ei. În loc să lanseze variante noi zilnic, Google alege să facă o mare schimbare.

La începutul anilor 2020, Google lansează penultimul sistem de operare, un mega sistem de operare care poate rula pe telefoane inteligente, difuzoare inteligente, laptopuri, tablete și electrocasnice cu funcționalități legate de internet. Acesta este doar începutul. În final, Google urmărește să crească și să îmbunătățească

acest sistem de operare astfel încât să devină infrastructura invizibilă care ne ajută în viața de zi cu zi, operând interfețele, căștile și ochelarii inteligenți, mașinile și chiar anumite componente ale orașelor. Acest sistem este complet integrat cu DDP-urile și aduce îmbunătățiri semnificative tuturor utilizatorilor săi.

Mega sistemul de operare al Google este lansat într-o perioadă nefastă pentru Apple, care a devenit prima companie americană care a atins o capitalizare de piață de o mie de miliarde de dolari, dar care înregistrează un declin constant al vânzărilor de telefoane iPhone pe fondul unui val de noi aparate conectate, precum căștile sau ceasurile inteligente. În ciuda multelor sale reușite, Amazon (a doua companie americană care a atins capitalizarea de o mie de miliarde de dolari) nu a mai avut o reușită în domeniul electronicelor de la difuzorul inteligent Echo încoace. Într-o mutare surpriză, Apple și Amazon încheie în 2025 un parteneriat pentru a dezvolta un sistem de operare care să ruleze pe aparatele produse de ambele companii. Sistemul de operare rezultat-Applezon- constituie o amenințare formidabilă la adresa Google. Pentru consumatori, această evoluție instaurează un model bazat pe două sisteme de operare și dă startul unei consolidări masive și rapide a ecosistemului inteligenței artificiale.

Facebook decide că trebuie să încheie un parteneriat similar; numărul utilizatorilor activi este în scădere, iar aceștia nu mai consideră rețeaua de socializare ca fiind indispensabilă. Facebook încearcă să se apropie de Applezon, care nu se arată interesați. Microsoft și IBM continuă să se axeze pe serviciile oferite companiilor.

China și noii săi parteneri diplomatici folosesc împreună tehnologiile BAT, în timp ce restul lumii folosește fie mega sistemul

de operare al Google, fie Applezon, ambele sisteme folosind și dând valoare DDP-urilor noastre. Piața ne oferă alegeri limitate. Există doar câteva modele de telefoane (iar ochelarii și brățările inteligente vor înlocui în curând telefoanele mobile) dar și pentru toate aparatele din casele noastre: difuzoare, computere, televizoare, electrocasnice sau imprimante. E mai ușor să alegem o singură marcă, astfel există case Google sau case Applezon. Din punct de vedere tehnic, ne-am putea transfera DDP-urile la alți furnizori, însă nu deținem personal datele din DDP-uri și nici DDP-urile însele. Nu beneficiem de o transparență totală- nu știm în bună măsură ce fac Google și Applezon cu DDP-urile noastre pentru că prin felul în care și-au creat sistemele au urmărit protejarea proprietății intelectuale.

Pentru a evita procesele anti-trust, ni se spune că putem porta oricând DDP-urile între sistemele de operare. În practică e însă aproape imposibil să faceți tranziția. Vă aduceți aminte cât de greu era acum mulți ani când încercați să treceți de la iOS la Android și descopereați că multe date și setări importante se pierdeau pentru totdeauna, că nu se păstra activitatea din cadrul aplicațiilor, multe dintre ele nici măcar nu funcționau (și nu puteați obține rambursarea banilor plătiți pentru ele) și că toate locurile în care vă depozitați înainte pozele și videourile nu mai puteau fi accesate cu ușurință. Acum că DDP-ul vostru este folosit și de terți-școli, spitale sau linii aeriene, trecerea de la Google la Applezon sau viceversa este mult mai dificilă.

Există mulți consultanți în informatică specializați în portarea DDP-urilor de la un furnizor la celălalt, însă procesul durează câteva zile, este costisitor și imperfect. Majoritatea oamenilor aleg

cu reticență să rămână unde sunt, chiar dacă rezultatele nu sunt optime.

Google și parteneriatul Amazon-Apple se confruntă cu procese anti-trust atât în Statele Unite cât și în Europa. În timp ce procesele se târăsc în instanță, toate datele sunt atât de amestecate încât separarea sau deschiderea sistemelor DDP și ale inteligenței artificiale ar genera riscuri mai mari decât recompensele potențiale. Ca urmare, se ia decizia ca cele două companii să fii amendate cu sume considerabile- fondurile urmează să fie folosite pentru dezvoltarea unor noi companii. Cu toții sunt însă de acord că sistemul cu doi furnizori de sisteme de operare trebuie să continue.

În timp ce inteligența artificială se maturizează și trece de la aplicații limitate la mașini capabile de inteligență generală, nu avem de ales și trebuie să trăim cu tăieturile cu hârtia provocate de inteligența artificială. Varianta modernă a comunismului chinez- socialism amestecat cu nuanțe de capitalism- se extinde, permițându-i lui Xi Jinping să îndeplinească promisiunile legate de noua ordine mondială. Țările care se opun stilului autocratic chinezesc, suprimării presei libere și libertăților religioase și perspectivelor negative asupra anumitor etnii, orientări sexuale sau de gen nu dețin suficientă putere. Ele nu au de ales și trebuie să colaboreze cu China, în termenii impuși de aceasta.

Ni s-a promis că inteligența artificială ne va aduce libertatea, eliminând sarcinile plictisitoare și munca repetitivă. În schimb, libertatea noastră de a alege este restricționată de o manieră pe care nimeni nu și-a imaginat-o.

2029: Neajutorarea învățată

Regimul binar de sisteme de operare a generat o concurență acerbă între membrii triburilor inteligenței artificiale care nu au anticipat problemele mari de interoperabilitate. Se pare că pe lângă aparate, nici *oamenii* nu sunt interoperabili în acest regim. Lejeritatea odinioară emblematică a Silicon Valley- ingineri, directori de operațiuni sau designeri ai experienței utilizatorilor obișnuiau să treacă de la o companie la alta fără niciun simț adevărat de apartenență- a dispărut de mult timp. În loc să ne aducă împreună, inteligența artificială a reușit efectiv și eficient să ne izoleze. Și pentru guvernul Statelor Unite reprezintă o problemă, fiind forța să aleagă o structură. (Ca majoritatea celorlalte guverne și cel american a ales Applezon în defavoarea Google, pentru că Applezon oferea prețuri mai mici și rechizite la preț redus).

În toată lumea se vorbește despre „neajutorarea învățată" în epoca inteligenței artificiale. Părem că nu putem funcționa fără diversele sisteme automate, care ne sugestionează în mod constant cu feedback negativ sau pozitiv. Încercăm să dăm vina pe cei Nouă Titani, dar de fapt doar noi suntem de vină.

Pentru membrii generației Y, lovitura a fost deosebit de grea, deoarece ei erau însetați să primească feedback și laude când erau copii și ei au iubit la început diversele sisteme ale inteligenței artificiale, însă între timp au căpătat un tic psihologic care e greu de eliminat. Când se termină bateria unei periuțe de dinți, un membru al acestei generații (care are acum peste 40 de ani) trebuie să-și perieze dinții cu o periuță obișnuită, care nu oferă niciun feedback. Neprimind niciun feedback de confirmare, persoana nu primește dopa-

mina la care se așteaptă, simțindu-se anxioasă și tristă. Fenomenul nu se întâlnește doar la membrii acestei generații. O senzație de neliniște ne bântuie pe toți. Investim în lucruri redundante, ne cumpărăm obiecte analogice de rezervă (cum ar fi periuțele de dinți din plastic, căștile vechi sau ochelarii Warby Parker) pentru a le completa pe cele care conțin inteligență artificială. Ne-am pierdut încrederea în fostul bun simț și în aptitudinile de bază necesare vieții.

Standardele rivale ale Google și Applezon ne aduc aminte de ce înseamnă să călătorești în străinătate și toate necazurile cu prize de format și tensiune diferite. Cei care călătoresc în mod regulat acordă prioritate alegerilor în funcție de sistemul de operare și nu de programele de fidelitate, alegând un hotel Applezon sau o linie aeriană Google. Companiile constată că e mai ușor să aleagă doar unul dintre cele două sisteme. Încet dar sigur, suntem sugestionați să alegem una dintre tabere. Cei care preferă Applezon au dificultăți în a trăi cu cei care preferă Google, pentru că DDP-urile și aparatele lor nu sunt compatibile- chiar dacă personalitățile lor sunt.

Anul 2019 a marcat începutul sfârșitului telefoanelor inteligente, motiv pentru care acum purtăm accesorii conectate și nu le mai ținem în buzunare sau în genți. După o perioadă de progrese rapide, noile telefoane care rulau iOS sau Android ofereau doar îmbunătățire minore ale sistemelor, în vreme ce pentru telefoane nu existau îmbunătățiri majore în afara fotografiei. Entuziasmul din jurul fiecării lansări a unor noi modele iPhone s-a pierdut. Nici măcar lansarea unui telefon Samsung cu ecran flexibil nu a fost suficientă pentru a convinge consumatorii să-și aban-

doneze modelele mai vechi. În loc să stea la coadă după un an sau doi pentru a achiziționa cel mai nou telefon, consumatorii au ales să cheltuie acei bani pe noi aparate: căști Bluetooth fără fir cu senzori biometrici, brățări care vă permit să înregistrați clipuri video și să efectuați apeluri video, și ochelari inteligenți care ne oferă un flux aparent infinit de informații. Applezon a reușit să lanseze prima pe piață ochelarii Applezon Vision- și nu a fost nicio surpriză. Apple și Amazon aveau un istoric lung și plin de reușite în promovarea noilor tehnologii și modelarea preferințele consumatorilor. (Eșecul comercial al Google Glass continuă să fie dureros pentru unii angajați ai companiei, chiar dacă tehnologia era revoluționară.) În zilele noastre majoritatea oamenilor poarte ochelari și căști inteligente toată ziua, cu o brățară sau un inel cu care pot efectua înregistrări video.

Se pare că ochelarii erau de neevitat. După două decenii în care ne-am holbat în ecrane, ochii noștri nu mai pot face ajustările necesare și majoritatea dintre noi avem o vedere încețoșată la distanță și avem nevoie de ochelari de vedere de la vârste mai fragede. Ca majoritatea oamenilor și voi aveți nevoie de ochelari, creând o piață pentru ochelarii inteligenți pe care unii analiști o credeau imposibilă. Ochelarii, împreună cu căștile fără fir, o brățară inteligentă și o tabletă ușoară, reprezintă principalul mijloc de comunicare. Vă uitați la videouri prin ei, iar pentru a efectua un apel video folosiți camera din brățara inteligentă. În general, vorbiți mai mult decât tastați. Algoritmi speciali pentru procesare în spațiu, vizualizare digitală și recunoaștere vocală stau la baza majorității datelor pe care le vedeți și le înregistrați prin intermediul acestor accesorii inteligente.

Applezon și Google v-au sugestionat să închiriați- și nu să dețineți- aceste echipamente, iar accesul la DDP face parte din abonament. Nu e nimic necurat în acest model al abonamentului, decizia fiind luată din considerente de ordin practic legate de ciclul de viață al produselor. Rata evoluției în domeniul inteligenței artificiale accelerează în fiecare an, iar de vreme ce valoarea datelor noastre este semnificativ mai mare decât marja de profit a ochelarilor, brățărilor și căștilor inteligente, obiectivul este de a ne menține pe toți conectați la sistemul. Tehnologia se vinde în pierdere, fiind compensată de un abonament lunar accesibil. Pe baza acestui abonament obțineți și accesul la DDP, care este evaluat în funcție de permisiunile acordate companiilor. Abonamentele cele mai ieftine oferă și anonimitatea cea mai scăzută, prin urmare clienții care aleg această variantă acordă libertatea Google și Applezon să folosească datele lor personale după bunul plac, fie că e vorba de publicitate sau de simularea unor experimente medicale. Persoanele mai întărite pot adăuga „prime de permisiune" pachetelor DDP, dar acestea se obțin cu mare dificultate și costă foarte mult. În 2029, există comunități ale elitelor, cu pază sporită și departe de ochii lumii-acestea sunt însă comunități digitale, păzite de algoritmi și care protejează datele persoanelor bogate de ochii curioși ai oamenilor și companiilor obișnuite.

Ca mulți alții, ați fost victime ale unor așa-zise „atacuri ale papagalilor", o nouă variantă de înșelătorii digitale, iar guvernele din lumea întreagă au fost luate complet prin surprindere. Se pare că programele ilegale vă pot infecta și DDP-ul, imitând ca un pa-

pagal vocea voastră către toți cunoscuții. Unele inteligențe artificiale de tip papagal sunt atât de bine înrădăcinate în dosarele voastre și în viața voastră digitală încât pot imita convingător nu doar vocea, ritmul, tonul și vocabularul vostru, dar o pot face cu date reale din viața voastră. Aceste programe sunt folosite pentru a trimite mesaje vocale false pe care mulți părinți sau soți le cred adevărate. Din nefericire, aceste programe de inteligență artificială papagal provoacă probleme companiilor care oferă servicii matrimoniale pe internet. Escrocii fură identitatea altora și o folosesc pentru a atrage oamenii folosind interacțiuni hiper-realiste.

Neajutorarea învățată, noile bariere economice și senzația că sinele nostru real nu poate concura cu varianta îmbunătățită prin inteligența artificială ne provoacă multe amărăciuni. Căutăm consolarea prin interfețele creier-mașină, legături cu capacitate mare care transferă date între creier și un calculator. Deși Facebook și Elon Musk au anunțat acum un deceniu că lucrează la un aparat special care ne-ar da super-puteri telepatice, Baidu a fost prima companie care a lansat „bentița neurostimulatoare." Ascunsă într-o șapcă, ea poate citi și monitoriza undele cerebrale și transmite un feedback pentru a crește concentrarea, a crea o senzație de fericire și mulțumire, sau a vă face să credeți că aveți multă energie. Nu a fost de mirare că o companie a grupului BAT a reușit să dezvolte prima o astfel de interfață. Companiile farmaceutice au făcut lobby pe lângă autorități, în speranța că acestea vor bloca autorizarea bentițelor neurostimulatoare și a unor interfețe viitoare între creier și mașină. Considerând că Baidu e o amenințare,

Google și Applezon au intervenit, lansând propriile produse, ceea ce a făcut ca DDP-urile noastre să conțină și mai multe date.

Sâcâiala e noua sugestionare, iar Google și Applezon vă hărțuiesc neintenționat pentru a avea o sănătate mai bună. Brățara, căștile și ochelarii inteligenți vă comunică în mod constant mesaje de reamintire. Nu apucați să mâncați o linguriță de tort, pentru că din clipa în care vă uitați către desert, inteligența artificială recunoaște ce urmează să mâncați, compară mâncarea cu ritmul metabolic și starea de sănătate actuală și vă trimite o notificare de avertizare către brățară sau ochelari. La restaurant, vi se sugestionează să alegeți mâncărurile care se potrivesc cu nevoile biologice actuale: alimente cu conținut ridicat de potasiu sau omega-3, sau alimente cu conținut scăzut de carbohidrați sau sare. Dacă alegeți inteligent, sunteți recompensați și primiți mesaje de încurajare.

Nu aveți nicio variantă reală de a vă deconecta de la inteligența artificială sâcâitoare, deoarece DDP-ul este legat de prima de asigurare de viață, calculată pe baza angajamentului de a duce o viață sănătoasă. Dacă săriți peste o sesiune recomandată de antrenament, vă puteți aștepta să fiți sâcâit toată ziua. Dacă mâncați o prăjitură în plus, fapta e consemnată la dosar. Sistemul nu a fost creat să se comporte în acest fel, însă algoritmilor li s-a oferit un scop și au fost învățați să optimizeze fără întrerupere diferitele aspecte ale vieții cotidiene. Nu au fost programați cu un punct de final sau o dată de încetare.

Când s-a instaurat regimul cu două sisteme de operare pentru DDP-urile noastre, mulți dintre furnizorii de arhive medicale au

trebuit să aleagă un partener. Unii membri ai Mafiei G au obținut astfel datele de care au avut nevoie în urmă cu mai mulți ani, creând- întrucâtva întâmplător-noul sistem sanitar american. IBM Watson Health deținea tehnologia complexă (unii ar putea susține chiar superioară) dar avea și un istoric de disfuncționalități organizaționale de două decenii. La 15 ani după ce Google a lansat Calico, propriul proiect de sănătate, acesta nu a reușit să scoată pe piață produse viabile din punct de vedere comercial, astfel încât un parteneriat strategic părea soluția: Watson Calico. Din partea Google a fost o mișcare inspirată deoarece atât Amazon cât și Apple plănuiau propriile inițiative în domeniul asigurărilor și farmaceuticelor. Amazon experimentase cu noi modele de asigurare și livrare a medicamentelor printr-un parteneriat cu Berkshire Hathaway și J.P. Morgan, în timp ce Apple și-a folosit lanțul de magazine și modelul Genius Bar pentru a lansa o nouă marcă de clinici la minut pe toată coasta de vest a Statelor Unite. Parteneriatul dintre Google și IBM a forțat o a doua colaborare de tip Applezon, între platforma electronică a farmaciilor Amazon și clinicile la minut ale Apple. Ca urmare a acestor consolidări de pe piață, spitalele din America fac astăzi parte fie din sistemul Watson-Calico, fie din Applezon Health Marile companii de asigurări- Kaiser Permanente, LifePoint Health, Trinity Health, NewYork-Presbyterian Healthcare System sunt înscrise fie în rețeaua Watson-Calico, fie în Applezon Health.

Aceste colaborări s-au dovedit o soluție extraordinară la problema datelor. Google, IBM și Applezon au acum acces nelimitat chiar și la datele voastre biologice-iar voi primiți la schimb acces la diagnostice gratuite sau cu cost scăzut. Analizele nu se efectuează

doar când sunteți bolnavi. Sunteți analizați pentru orice, lucru care are un efect benefic asupra nivelului de sănătate și fericire. Dacă întrebați orice american ce temperatură are corpul lui, veți primi un răspuns personalizat, și nu cele 37,5 grade Celsius standard.

Deși avem în sfârșit acces la asistență sanitară accesibilă, americanii trebuie acum să suporte niște defecte bizare care s-au dovedit a fi trăsături ale sistemului și nu erori. Ambulanțele mai vechi sau asistentele medicale din școli sau din taberele de vară nu pot accesa întotdeauna DDP-ul pacientului dacă nu au cele mai recente actualizări de sistem. DDP-urile sistemelor rivale de sănătate pot fi citite din punct de vedere tehnic atât de Applezon Health cât și de Watson-Calico, însă adesea lipsesc multe date contextuale utile. În special în comunitățile mai mici sau cele rurale, doctorii constată că trebuie să-și aducă aminte ce au învățat în facultate dacă o persoană care locuiește într-o casă Applezon se prezintă la o clinică Watson-Calico, sau invers. Pe măsură ce doctorii instruiți în metodele mai vechi, tradiționale se pensionează, există din ce în ce mai puțini doctori tineri care au cunoștințele și experiența necesară pentru a putea consulta pacienții incompatibil. Iată încă un exemplu de neajutorare învățată, dar cu consecințe dintre cele mai grave.

Inteligența artificială a afectat și alte domenii în mod bizar. În 2002, cercetătorii de la centrul Open Infrastructure for Network Computing al universității Berkeley și-au dat seama că dacă am putea permite ca calculatoarele noastre să lucreze pentru alții în timp ce dormim, s-ar putea simula puterea unui super-calculator, iar acea putere ar putea fi folosită în scopuri științifice. Primele ex-

perimente s-au dovedit utile, iar sute de mii de oameni au donat timpul mort de procesare unor proiecte nobile din lume, cum ar fi Quake-Catcher Network, o rețea care caută activitatea seismică, sau SETI@home, care caută viața extraterestră în univers. În 2018, unii antreprenori inteligenți și-au dat seama cum aceste rețele ar putea fi reconfigurate pentru varianta actualizată a unei economii bazate pe sarcini de scurtă durată. În loc să conducă pentru Uber sau Lyft, persoanele pot instala aplicații pe cont propriu prin care să câștige bani din timpul lor liber. Cele mai recente aplicații permit companiilor să folosească aparatele noastre, în schimbul unor credite sau sume de bani pe care le putem cheltui în alte scopuri.

La fel ca la începutul serviciilor de ride-sharing, mulți oameni au părăsit forța de muncă tradițională pentru a lua parte la această variantă nouă a economiei bazată pe sarcini de scurtă durată. Și-au dat demisia, și au încercat să câștige bani închiriind accesul la aparatele lor. Acest fenomen a provocat o presiune semnificativă asupra furnizorilor de curent electric și internet, care nu au mai putut face față cererii. În zilele noastre, supraîncărcarea rețelei și penele de curent sunt des întâlnite, iar din cauza faptului că oamenii au tendința de a rula aceste programe în timpul somnului, ei nu-și dau seama că au pierdut un venit potențial până dimineața.

Cei care au rămas în forța de muncă tradițională au început să folosească inteligența artificială pentru a-și actualiza CV-urile și scrisorile de intenție, provocând o altă problemă. Problemele care ar fi putut elimina anumiți candidați sunt acum mai puțin evidente- toți par acum că au un avantaj competitiv. Sistemele de inteligență artificială sunt folosite pentru a tria candidații, dar managerii nu mai pot alege pentru că toți candidații par la fel de buni.

Ei recurg așadar la un comportament care să le dea o senzație de confort: bărbații albi au tendința să angajeze bărbați albi, pentru că sunt paralizați de tirania alegerii.

În majoritatea companiilor mari, ierarhia anterioară a fost redusă la două niveluri de muncitori: angajați cu aptitudini și management superior. Angajații cu aptitudini lucrează împreună cu sisteme de inteligență artificială și raportează unor superiori din inteligența artificială, pentru că întreaga pătură a managementului de mijloc a fost eliminată. La muncă, sistemele de supervizare monitorizează productivitatea, vă urmăresc în birou, notează cu cine socializați și înregistrează nivelul vostru de fericire, anxietate, stres și mulțumire. Sistemele sunt o întruchipare a acelor postere motivaționale oribile, care vă aduc aminte că „sunteți mai curajoși decât credeți" și „sunteți mai puternici decât pretextele".

Guvernele au fost luate prin surprindere de eliminarea la scară largă a managementului de mijloc în industriile bazate pe cunoștințe -cum ar fi dreptul sau finanțele- pentru că s-au concentrat exclusiv pe munca sau ocupațiile cu aptitudini reduse cum ar fi șoferii, fermierii sau muncitorii. Domeniile creative sunt lovite la fel de greu de o nouă ramură a inteligenței artificiale: creativitatea mașinilor. Graficienii, arhitecții, copywriterii și dezvoltatorii web au devenit inutili pentru că rețelele antagonice generative și sistemele de inteligență artificiale mai recente s-au dovedit a fi remarcabil de productive și sigure. În același timp, inteligența artificială a conferit super-puteri anumitor posturi- directorii generali, financiari sau tehnici. S-a căscat o prăpastie semnificativă, și vârful

organizațiilor concentrează din ce în ce mai multe averi. Constatăm începutul unui sistem de caste în epoca digitală.

Încă un defect: contaminarea informațiilor. În urmă cu un deceniu, multe procese și reglementări internaționale au provocat fracturarea internetului. În locul unei singure rețele mondiale, am ajuns să avem mai multe rețele paralele, în care regulile în domeniul digital depind de legile naționale și de restricțiile geografice. Procesul nu s-a produs imediat. Când internetul a trecut de la guvern și mediul academic înspre mediu privat, în anii '90, i-am permis să se răspândească liber, în loc să impunem reglementări, precum în domeniul utilităților sau cel financiar. La vremea aceea, legiuitorii nu s-au gândit prea mult la felul în care ar putea fi folosite ulterior toate datele pe care le generăm pe internet. Acum a devenit imposibil de respectat fiecare lege diferită a fiecărei țări, iar filtrele anterioare de conținut s-au extins pentru a corespunde granițelor geografice. Astfel, s-a contribuit la promovarea și propagarea știrilor false. Deoarece actorii malefici folosesc algoritmii de generare și pentru că primim știri diferite în funcție de regiune, nu știm în cine sau în ce să avem încredere. Toate marile surse de știri din lume au fost păcălite cel puțin o dată, deoarece până și jurnaliștii cu experiență întâmpină dificultăți în a verifica atât videoclipuri cu lideri globali dar și cu oameni obișnuiți. E aproape imposibil să ne dăm seama dacă vocile și chipurile din clipurile pe care le urmărim sunt simulate sau adevărate.

O altă problemă: apariția unui val de criminalitate bazată pe inteligența artificială pe care nimeni nu l-a anticipat. Programe cu aplicații reduse dar puternice au început să provoace probleme pe internet. Acestea se ocupă cu achiziții ilegale: genți de marcă contrafăcute, droguri sau medicamente produse din animale braconate (cum ar fi coarnele de rinocer sau fildeșii de elefant). Aceste programe ne spionează pe rețelele de socializare, citesc știrile și infiltrează piețele financiare provocând reprize bruște de panică. În spațiul public, aceste programe calomniază, cu intenția de a defăima caracterul și reputația oamenilor. Începem să ne facem probleme că inteligența artificială ne-ar putea accesa DDP-urile, modifica datele biometrice și falsifica atât propriile date, cât și cele pe care le-am moștenit. O parte a acestor nelegiuiri sunt provocate intenționat de către varianta modernă a mafiei: o rețea generalizată de criminalitate organizată dificil de depistat și ținut sub control. Unele dintre inteligențele artificiale malefice sunt însă rodul întâmplării: ele au evoluat pur și simplu și au început să se comporte de o manieră pe care nimeni nu o plănuise.

Problemele se extind și la roboții fizici. Roboții de securitate, dotați cu camere inteligente și programe analitice predictive, au tendința să urmărească în mod regulat oamenii de culoare. Roboții de securitate nu sunt dotați cu arme, dar emit comenzi cu voce foarte tare și alarme cu sunete ascuțite dacă bănuiesc că se petrece vreo nelegiuire. În interiorul clădirilor de birouri, al hotelurilor, gărilor și aeroporturilor, oamenii de culoare sunt hărțuiți și umiliți în mod constant pentru că un robot de securitate i-a etichetat ca fiind suspecți.

Mafia G nu are o relație ușoară cu forțele de ordine americane, care vor să acceseze DDP-urile. În loc să colaboreze, autoritățile amenință cu procese și încearcă să oblige Mafia G să partajeze datele, iar companiile nu au nicio obligație -de ordin juridic sau altminteri- de a ceda acestor solicitări. Deși nimeni nu vrea să declare acest lucru în public, se pare că autoritățile americane speră să poată replica anumite părți ale monitorizării algoritmice și a sistemului de credite sociale din China. Temându-se de represaliile consumatorilor, Mafia G continuă să refuze accesul la sistemele lor.

Am vorbit de peste un deceniu despre impactul filosofic și etic al procesului decizional în cadrul forțelor de ordine, însă fără a se ajunge la vreun standard, vreo normă sau vreo reglementare. Avem acum un șir aparent infinit de infracțiuni comise cu ajutorul inteligenței artificiale, fără a avea însă metode de pedepsire. Nu există închisori pentru inteligențe artificiale sau roboți. Legile care definesc infracțiunile nu se aplică și tehnologiilor pe care le-am creat.

Confuzia și dezamăgirea noastră se pliază de minune pe strategia Chinei, care nu e doar un rival care tinde spre egalitatea cu Statele Unite, ci un concurent direct și o amenințare de ordin militar. China a petrecut decenii furând designul echipamentelor americane și strategiile de apărare, iar această tactică a adus beneficii. Președintele Xi consolidează și mai mult puterea armatei chineze, concentrată mai mult pe programare decât pe luptă fizică. Spre exemplu, spectacolele luminoase chinezești prezentate cu ocazia diferitelor evenimente – un festival din 2017 cu „lampioane cu drone", o extravaganță de vară din 2018 cu artificii cu drone-s-au dovedit a fi teste pentru supravegherea maselor. Armata Chi-

nei folosește acum drone puternice bazate pe inteligența artificială care vânează în haite, pe uscat și pe mare.

Prin intermediul puterii economice, diplomației directe și puterii militare, China a deschis o nouă epocă colonială, colonizând cu succes Zambia, Tanzania, Republica Democrată Congo, Kenya, Somalia, Etiopia, Eritreea și Sudanul. China construiește infrastructura-și implementează sistemul de credite sociale- și extrage resurse naturale esențiale pentru a-și sugruma concurența și a-și alimenta clasa mijlocie aflată în expansiune. China controlează acum 75% din rezervele mondiale de litiu, o materie primă esențială pentru baterii. China a distrus și pădurile de santal și a condus la dispariția copacului Mukula, o specie cu creștere lentă din Africa centrală care fusese folosită o perioadă pentru a fabrica mese și scaune sculptate colorate în roșu.

Nicio putere economică-nici Statele Unite, Japonia, Coreea de Sud sau Uniunea Europeană- nu are suficientă influență politică sau economică pentru a împiedica China să-și extindă zonele de influență economică specială în Marea Chinei de Sud, Marea Chinei de Est și Marea Galbenă. Aproape jumătate din comerțul global trebuie să tranziteze una din aceste zone, și fiecare navă care tranzitează trebuie să plătească o taxă considerabilă guvernului chinez.

Unii analiști spun că Beijingul a ratat ținta de a deveni principala putere în domeniul inteligenței artificiale în 2025, chiar dacă a preluat controlul asupra anumitor resurse din lumea fizică. Acești analiști nu văd însă imaginea de ansamblu. După ani de contracte de transferuri obligatorii de tehnologie, practici de restrângere a concurenței, și investiții considerabile chineze în companiile europene și americane, strategia Chinei s-a dovedit extrem de

reuşită. China domină în domeniul tehnologiilor avansate, inclusiv în ce priveşte robotica, sursele noi de energie, genetica şi aviaţia- iar toate aceste domenii se folosesc de inteligenţa artificială şi sunt îmbunătăţite cu ajutorul acesteia. Nu există date oficiale, dar luând în calcul laboratoarele de cercetare de stat, parteneriatele cu Baidu, Alibaba, şi Tencent şi partenerii din proiectul Noul Drum al Mătăsii, experţii estimează că China a reuşit să sporească valoarea ecosistemului inteligenţei artificiale cu peste 500 de miliarde de yuan (aproximativ 73 de miliarde de dolari) doar într-un deceniu.

2049: Şi au rămas doar cinci

Pe măsură ce timpul trece şi se înregistrează progrese în domeniul inteligenţei artificiale generale, constelaţia celor Nouă Titani s-a modificat de o manieră atât profundă cât şi problematică. Grupul BAT din China e acum mai puternic ca niciodată, şi continuă să primească ordine de la Beijing. Din cei şase membri originali ai Mafiei G, au rămas acum doar cinci, din cauza parteneriatelor strategice şi alianţelor: Amazon-Apple şi Google IBM sunt companiile cu cea mai mare putere. Microsoft oferă în prezent doar asistenţă tehnică pentru sistem şi servicii depăşite.

Poate cea mai surprinzătoare evoluţie este însă cea a Facebook. Nici scandalul Cambridge Analytica şi nici măcar dezvăluirile despre amestecul Rusiei în alegerile din Statele Unite nu au condus la prăbuşirea Facebook. Nu a fost nici oboseala pe care o resimţeam cu toţii, în timp ce fluxurile se umpleau din ce în ce mai mult cu ură, teorii ale conspiraţiei politice şi demagogie. Modelul de afaceri al Facebook s-a dovedit a nu fi sustenabil pe termen lung. Odată

ce utilizatorii au început să nu mai folosească rețeaua și companiile de publicitate nu au mai cheltuit bani pe platformă, Facebook s-a găsit în postura de a nu avea venituri din mai multe surse. Până în 2035, compania a început să aibă probleme financiare grave. Acționarii voiau să vândă, directorii fondurilor mutuale s-au speriat, și piața s-a întors împotriva companiei, care a fost vândută pe bucăți. Toți cei ale căror date erau blocate în interiorul rețelei- majoritatea americanilor- sunt acum foarte îngrijorați, pentru că datele noastre au fost cumpărate pe ascuns de un consorțiu. Ancheta e în desfășurare, dar se pare că acest consorțiu e de fapt o companie controlată de China. E probabil ca acum să facem cu toții parte din sistemul chinezesc de credite sociale și că suntem cu toții urmăriți.

Și voi, la fel ca ceilalți americani, ați învățat să trăiți cu o anxietate constantă, de nivel scăzut. Nivelul național de nelinișteste adesea comparat cu amenințarea războiului nuclear din anii 1960 și 1980. De data aceasta însă, obiectul concret al fricii este necunoscut. Nu știm dacă DDP-ul este în siguranță, sau la ce date are acces China. Sunteți nesiguri cât de infiltrați sunt hackerii chinezi în infrastructura țării noastre. Vă treziți adesea dimineața devreme, întrebându-vă ce știe China despre voi, despre podurile pe care mergeți spre muncă, despre țevile de gaz care vă intră în casă, și ce plănuiesc să facă cu toate aceste informații.

Ceea ce nu am putut anticipa a fost un spectru de IAG, construite pentru diferite scopuri și sarcini, care sunt atât puternice dar și indiferente la valorile umane. Dacă privim în urmă, am fost extrem de naivi. Pe măsură ce Amazon, Apple, Google și IBM au încheiat parteneriate, au ales o tabără sau alta și au crescut, aceste companii nu și-au setat și standarde globale. Cu decenii în urmă,

oamenii cumpărau aplicații sau jocuri din magazinul Google Play, și pentru că aproape oricui îi era la îndemână să lanseze și să vândă o aplicație, calitatea diferea semnificativ. Existau mult prea multe aplicații care consumau bateria rapid, jocuri care colectau și partajau informații personale și reclame enervante care stricau experiența pe mobil. Aceleași fenomene se manifestă acum în cazul IAG, însă cu consecințe mult mai grave. Unele IAG se prefac că urmează protocoalele pe care le-au primit, apoi le suprascriu cu noi instrucțiuni. Unele IAG se îmbunătățesc singure, chiar dacă cei care le-au creat nu le-au programat să facă asta. Unele se înmulțesc singure, se subdivid în alte IAG și obțin resursele de care au nevoie pentru a-și îndeplini obiectivul, indiferent de impactul pe care l-ar putea avea aceste acțiuni asupra ecosistemului.

Pentru a contracara problema IAG care nu se comportă cum trebuie, cercetătorii Applezon și Google-IBM au implementat programe IAG dădacă -IAGD pe scurt- pentru a ține în frâu alte sisteme. IAGD au un set clar de protocoale:

- să investigheze și analizeze alte IAG, pentru a vedea dacă acestea încalcă obiectivele stabilite inițial
- să înregistreze în amănunt toate IAG care nu se comportă cum trebuie, dar și întregul lor istoric (de ex., cine le-a creat, când au fost modificate, și de către cine/ce)
- să găsească primul programator și să-l înștiințeze de faptul că IAG nu se comportă cum trebuie
- după o perioadă de grație (care depinde de gravitatea infracțiunilor IAG) să scoată din uz orice IAG
- să nu-și modifice niciodată propriile obiective

Este evident că atât Applezon cât și Google-IBM încercau să controleze un sistem care începea să scape de sub control, însă în afara ecosistemelor Applezon și Google IBM, IAGD nu au fost adoptate la scară largă. Folosind ca precedent hotărârile anti-trust împotriva Google și Microsoft, Parlamentul European a susținut că IAGD nu sunt nimic mai mult decât o tentativă disimulată a companiilor de a înăbuși antreprenorii și de a zdrobi concurența. UE a devenit prima care a interzis IAGD. Chiar și în timp ce cercetătorii le cereau autorităților să permită acestor IAG specializate să contribuie la limitarea unei probleme serioase și în curs de agravare, Congresul Statelor Unite a votat împotriva giganților din domeniul tehnologiei, interzicând folosirea IAGD. Aceste decizii lipsite de viziune nu au făcut decât să sădească o neîncredere a publicului în Applezon și Google-IBM, care ar fi putut fi altminteri niște păzitori de încredere ai DDP-urilor noastre.

Casa voastră a devenit un mare receptacul pentru marketing, care are un caracter constant și invaziv. Vedeți reclame personalizate oriunde există un ecran: în oglinzile inteligente din baie și dulap, în ecranele retractabile din buzunare, și chiar în geamurile inteligente pe care a trebuit să le instalați în case pentru a bloca radiațiile solare extreme. Vă simțiți incomod în propria casă- locul unde odinioară erați cel mai relaxați.

Această neîncredere a transformat sistemul de sănătate într-un monstru. Applezonn Health System și Watson Calico au înregistrat progrese deosebite atât în domeniul inteligenței artificiale cât și al medicinei. Ambele au preluat ideea de la un costum robotic controlat cu

mintea prezentat la Cupa Mondială din 2014. Neurobiologul Miguel Nicolelis de la Duke University și-a dat seama cum să îmbine mintea și mașina-iar studiile lui i-au inspirat și pe alții să scoată pe piață interfețe creier-mașină. În anumite birouri unde tehnologia este mai avansată, muncitorii sunt încurajați să poarte bentițe electronice și să-și conecteze mințile, împreună cu o IAG, pentru a rezolva probleme dificile. Nu toți oamenii se simt confortabile cu această variantă hipertehnologizată a inteligenței colective, deoarece interacțiunea presupune ca datele să treacă fie prin Applezon sau prin Watson Calico, care pot acum vedea la propriu ce ne trece prin cap.

Watson Calico a reușit, în parteneriat cu o universitate de top din New York, să ducă mai departe una dintre teoriile mai puțin cunoscute ale lui Turing legată de morfogeneză și inteligența artificială. Turing credea că un sistem de substanțe chimice reacționează probabil unele cu celelalte, iar acea reacție se transmite de la o celulă la alta, modificând o parte dintre aceste celule. S-a dovedit că Turing a avut dreptate. S-au folosit sistemele IAG pentru a descoperi diferite moduri prin care să se creeze ființe multicelulare, dând astfel naștere unor ființe umane modificate, pe care le denumim „himere om-animal."

Intenția originală era de a crea țesuturi omenești care să poată fi transplantate, așa că am folosit porcii și oile pentru a crește ficați, inimi și rinichi. Cercetătorii au dezvoltat și țesut neuronal. Domeniul părea promițător, până când ne-am dat seama că IAG era folosită pentru a dezvolta himere om-animal care aveau alte caracteristici cum ar fi porci implantați cu țesut cerebral omenesc care ajungeau la un IQ similar cu al unui om neinteligent și copii nou-născuți cu un simț al mirosului ca de câine. Ceea ce nimeni

nu a dezbătut (sau stabilit) sunt implicațiile pe care le presupun aceste atribute himerice, care sunt moștenibile. Ce se întâmplă atunci când un om modificat pentru a avea capacități extrasenzoriale are un copil cu o persoană care are la rândul ei modificări?

Cel mai îngrijorător este faptul că China a decis să le dea un nou scop interfețelor creier- mașină și IAG- ele fiind menite inițial să ajute bolnavii să-și recapete facultățile- în scopul obținerii unor avantaje strategice de ordin militar. Sistemele au fost folosite pentru a spori capacitățile cognitive ale soldaților, care luptă în mare măsură din interiorul unor bunkere subterane întunecoase. În Statele Unite și Uniunea Europeană, aceste experimente și aplicații ale tehnologiei încalcă legile privind etica.

Din cauza practicilor coloniale chineze, a expansiunii zonelor de influență economică și a folosirii lipsite de scrupule a IAG, începem să constatăm un declin real al civilizației occidentale și al idealurilor noastre democratice. Sănătatea economiei noastre este în pericol, iar indicatorii tradiționali precum numărul de construcții noi sau vânzările de alimente și ale magazinelor scad de la un trimestru la altul.

Chiar și Applezon și Google-IBM înregistrează într-un final o scădere a veniturilor, companiile fiind îngrijorate în privința viitorului. În timp ce lucrează pentru a moderniza DDP-urile pentru ca ele să poată lucra cu IAG paznici, ambele companii detectează o serie de date ciudate în sistemele de înregistrare. Există fragmente de cod care par să nu aibă sens, iar o parte din IAG care procesează și direcționează DDP-urile noastre se mișcă cu o viteză redusă. Într-un gest rar de colaborare, cele două companii parta-

jează rezultatele analizelor în speranța de a putea stabili problema. La birou și acasă, luminile se închid la întâmplare. Ochelarii noștri inteligenți au perioade intermitente în care nu funcționează. Sateliții noștri de comunicații ies de pe orbită.

Deși nu le putem auzi, știm că primele lovituri de tun au fost trase, și că China a declarat război Americii.

2069: Ocupația digitală a Statelor Unite ale Americii

Ne dăm seama că China și-a dezvoltat o generație de IAG care au capacități mult mai mari decât până în prezent. În absența unor IAGD care să păstorească IAG care o iau razna, China a putut construi și lansa un sistem îngrozitor prin care să controleze majoritatea oamenilor de pe Pământ. Dacă nu acceptăm pretențiile Chinei, aceasta ne va deconecta sistemele de comunicații. Dacă nu oferim Partidului Comunist Chinez accesul la datele noastre, acesta va închide infrastructura de bază, precum centralele electrice sau controlul traficului aerian.

Sunteți un cetățean în Statele Unite ale Americii sub ocupație digitală chineză. Transportul, banca, sistemul de sănătate, comutatoarele de lumini și frigiderele sunt toate la cheremul Chinei.

Ceea ce a început ca o incursiune colonială în Africa a generat apariția unui nou imperiu chinez, construit pe baza inteligenței artificiale. Omenirea se află în pragul unei SIA dezvoltate de către o țară care nu împărtășește valorile și idealurile noastre democratice.

CAPITOLUL ŞAPTE

DINASTIA RÉNGŌNG ZHÌNÉNG: SCENARIUL CATASTROFAL

„Aşa se sfârşeşte lumea, nu cu mare zgomot, ci cu un oftat."

—T.S. ELIOT

Până în 2023, am ignorat complet parcursul evolutiv al inteligenței artificiale. Am ratat toate semnalele de alarmă, le-am ignorat și nu ne-am imaginat în mod activ viitorul. I-am ajutat pe cei nouă Titani să concureze împotriva lor înșiși, ne-am dat frâu liber dorințelor de consum, cumpărând cele mai recente aparate, sărbătorind orice nouă ocazie de a ne înregistra vocile și chipurile, și transmițând aceste informații printr-o rețea care nea- absorbit în mod continuu datele. Am partajat videoclipuri stupide cu Alexa care nu poate avea o conversație corectă cu copiii noștri. Am cerut ca televizoarele să poată să ne scaneze fețele, fără să ne punem problema ce nevoie ar putea avea un televizor de datele noastre biometrice. De fiecare dată când Google lansa câte un nou proiect amuzant care ne mapa corpul în fotografii, chipul în tablouri, vocea pe a unor celebrități, amprentele noastre pe oameni din țări îndepărtate, și irisul pe strămoșii noștri, am luat parte cu interes, disperați să ținem pasul cu influencerii digitali și cele mai recente meme.

Triburile inteligenței artificiale spun că diversitatea contează. Aceasta e mantra lor. Ei repetă în mod constant această afirmație, la lansări și conferințe, la interviuri de angajare și la ședințele consiliilor de administrație, pe Twitter sau în articole. O spun în pliantele facultăților. O spun prin postere atrăgătoare atârnate în lifturi sau lipite pe holurile birourilor. Tribul inteligenței artificiale, în mare parte format din bărbați albi, este antrenat să recite această mantră în sălile de curs, în laboratoare și în spațiile de lucru. În loc să facă alegeri dificile și schimbări , aceștia repetă mantra și promit că schimbarea va veni curând. Iar mantra funcționează chiar în scopul în care a fost inventată: să elimine negativitatea din minte și să facă triburile inteligenței artificiale să se simtă mai bine în propria piele. Diverși guru ai triburilor inteligenței artificiale transmit mantra fiecărei noi grupe de discipoli, care trăiesc senzația unei mari reușite atunci când o repetă.

Mantra este repetată în bula comodă a triburilor inteligenței artificiale, care cred că ei promovează diversitatea în timp ce fac exact opusul. Ei promovează diversitatea de toate felurile -politică, religioasă, sexuală, de gen, rasială, etnică, de statut economic și vârstă- însă nu fac niciun demers serios pentru a fi mai incluzivi. În loc să vedem un spectru mai colorat de oameni și păreri în domeniul inteligenței artificiale prin poziții de vârf în mediul universitar, cercetare sau în Mafia G, nu vedem nicio schimbare.

Pe măsură ce perspectivele triburilor asupra lumii devin din ce în ce mai înguste, problemele pe care le-am amintit deja se înmulțesc. Accidentele și greșelile continuă să devină mai frec-

vente, de pildă sistemele de vizualizare digitală care identifică în mod eronat persoanele de culoare și le acuză de comiterea unor infracțiuni. Supravegherea se extinde devenind în același timp mai puțin evidentă. Linia de demarcație dintre datele personale și cele pe care le generăm la muncă se estompează, la fel și criteriile pe baza cărora se decide cine folosește datele noastre, și când. Transparența sistemelor de inteligență artificială dispare cu totul. (Nu că ar fi fost vreodată grozavă.)

Mafia G sunt singurii proprietari ai datelor voastre personale, care ajung să includă toate aspectele vieții: conținutul e-mailurilor, mesajele pe care le trimiteți copiilor, firmiturile digitale pe care le lăsați în urmă căutând scaunul perfect de birou, contururile unice ale degetelor și feței voastre, unde vă plimbați și cât de repede mergeți, peste cine dați la magazin, dacă aveți gripă și ce medicamente luați. Algoritmii iau decizii în numele vostru folosind toate datele voastre. Ei stabilesc dacă primiți sau nu o reducere când rezervați un zbor. Ei vă ajută sau vă împiedică să obțineți o slujbă, vă permit să vă cumpărați o casă sau o mașină, vă oferă parteneri pentru prima întâlnire și le comunică doctorilor dacă mințiți sau nu în privința consumului de alcool, tutun sau a exercițiilor fizice. Deoarece Google, Amazon, Aplpe, Facebook, Microsoft și IBM dețin aceste date- și pentru că iubim produsele lor, chiar dacă nu avem încredere deplină în aceste companii- nu vedem că faptul că aceste companii dețin controlul complet asupra DDP-urilor noastre reprezintă doar o variantă americană a sistemului de credite sociale din China.

Ne aflăm blocați într-un sistem de caste digitale, în care inteligența artificială face alegeri și judecăți bazate nu doar pe felul în care noi ne trăim viața, ci și pe DDP-urile părinților și rudelor

noastre. Averea nu mai contează. Statutul este stabilit de un „comportament exemplar" iar definiția „exemplarului" a fost dată acum mult timp de o minoritate de programatori care credeau că cursurile de yoga de la prânz, vizitele dese la chiropracticieni și o dietă organică ketogenică. Dacă nu faceți săptămânal o saună cu infraroșii, sistemul de inteligență artificială la care sunteți legat va înregistra abaterea în DDP. Acest act de răzvrătire nu vă va afecta doar pe voi, pentru că dosarul personal este conectat la toate persoanele pe care le cunoașteți și cu care sunteți înrudite. Nu puteți scăpa de păcatele celor din jur.

În viitorul apropiat, Amazon și IBM vor încerca să convingă guvernele Statelor Unite, Regatului Unit, Germania și Japonia să ofere accesul la o serie de date de sănătate ale cetățenilor. Apple, Google, Microsoft și Facebook vor întâmpina probleme și mai mari din cauza proceselor anti-trust anterioare. Experimentele Amazon și IBM se vor dovedi utile agențiilor guvernamentale, care vor oferi mai multe contracte profitabile pentru toată Mafia G.

În 2008, când multe țări ale lumii au intrat într-o criză financiară provocată de o bulă speculativă imobiliară, China s-a arătat dornică să cumpere fier, petrol și cupru din America Latină, protejând țările din regiune de daune economice mai mari. Când prețul petrolului a scăzut în 2011, China a fost dispusă să investească în America Latină și să salveze regiunea de la faliment. [179]

În 2013, China a efectuat exerciții militare comune pe coasta Braziliei-repetând manevra în 2014 pe coasta Chileului.[180] În 2015, ministerul apărării din China a găzduit un summit de 10 zile dedicat logisticii militare, cu invitați din 11 țări latino-americane, invitând apoi în anii următori ofițeri din aceste țări să participe la programele de dezvoltare a carierei în China.[181] În timp ce guvernul american bate în retragere și joacă un rol din ce în ce mai mic pe scena lumii, China pare axată pe expansiune, încheind înțelegeri cu țări din Asia de sud-est, Africa -și, desigur, din America Latină.

După un deceniu în care și-a construit temeinic relațiile în America Latină, China -și nu Statele Unite- este cea care furnizează echipament militar Venezuelei, Boliviei, Peruului și Argentinei, inclusiv avioane și arme.[182] China are un motiv pentru care să înființeze baze în tot hinterlandul Statelor Unite. În Patagonia, China a construit o antenă militară și o stație de control spațial, precum și un centru de urmărire a sateliților în nord-vestul Argentinei.[183] Toate aceste activități presupun folosirea inteligenței artificiale.

Politicienii nu sunt capabili să facă legătura între China, Statele Unite și inteligența artificială. Consolidarea puterii lui Xi Jinping, diferitele proiecte ale statului chinez, economia aflată în creștere galopantă și succesul grupului BAT reprezintă o forță de neoprit-chiar dacă invizibilă- și care trebuie confruntată. Nici Casa Albă și nici Congresul Statelor Unite nu văd că imixtiunile Chinei în toate aceste țări- Tanzania, Vietnam, Argentina, Bolivia- privesc atât rațiuni economice *cât și* de informații. Ei refuză să recunoască că China construiește un imperiu al secolului 21 bazat pe date, infrastructura inteligenței artificiale, geopolitică și economia globală. E o eroare gravă de judecată pe care o vom regreta cu toții ulterior.

Cetățenii chinezi învață să trăiască cu monitorizarea automată și consecințele încălcării regulilor. Criminalitatea se reduce, și neliniștea socială este înăbușită, iar pentru o perioadă păturile superioare și cele de mijloc își păstrează bunăstarea. Au acces la haine și genți de lux, mobilă de designer și mașini elegante pe care părinții sau bunicii lor nici nu și le-ar fi putut imagina. Se fac promisiuni că într-o bună zi sărăcia va fi eliminată din China. Pentru moment, se pare că renunțarea la intimitate, libertate religioasă, identitate sexuală și libertatea de expresie reprezintă un preț rezonabil de plătit în schimbul unui scor bun al creditelor sociale.

Liderii politici americani nu se interesează suficient în privința a ce este, ce nu este și de ce este importantă inteligența artificială. În afara discuțiilor obișnuite legate de perturbarea productivității și a șomajului, conducătorii de la Washington nu fac niciun efort de a comunica serios cu Mafia G în privința celorlalte aspecte legate de inteligența artificială, cum ar fi siguranța națională, echilibrul geopolitic, riscurile și oportunitățile inteligenței artificiale generale, sau introducerea inteligenței artificiale în alte domenii (cum ar fi genetica, agricultura și educația).

În absența unei direcții strategice referitoare la inteligența artificială din partea Casei Albe-de fapt, în prezența unei atitudini ostile față de știință și tehnologie- autoritățile de la Washington se concentrează mai mult asupra aspectelor relevante prin prisma următorului ciclu electoral și a veștilor care vor fi bine primite de showurile politice de duminică.

Nici Mafia G și nici conducătorii acestor companii nu iau măsuri care să pericliteze deliberat democrația. Protejarea statului de superputere globală dominantă al Americii și asigurarea conser-

vării idealurilor democratice nu fac însă parte din valorile de bază ale acestor companii. La începutul anilor 2010, fostul director al Google Eric Schmidt lucra intens și admirabil pentru a spori gradul de pregătire al armatei și guvernului Statelor Unite în epoca inteligenței artificiale. Nu era vorba de o tentativă de a obține contracte guvernamentale pentru Google. Schmidt era preocupat de siguranța națională și de capacitatea de reacție militară în această nouă eră tehnologică. Demersul era însă atât de neobișnuit încât companiile din Silicon Valley i-au pus în discuție motivele. În loc ca ceilalți conducători ai companiilor din Mafia G să-i urmeze exemplul, aceștia s-au arătat sceptici vizavi de ambițiile lui. Prin urmare, în afară de Schmidt, nici un alt conducător al Mafiei G nu s-a mai gândit la rolul pe care îl joacă inteligența artificială în planul Chinei de a uzurpa poziția dominantă a Americii.

Între Mafia G și agențiile guvernamentale sau militare nu există o colaborare strategică-cel puțin nu în afara unor contracte lucrative. Mafia G acceptă politicile complicate și demodate de achiziții ale armatei și guvernului, însă acest fapt nu contribuie la o evoluție a inteligenței artificiale în interesul național. Această acceptare nu face de fapt decât să releve diferențele culturale dintre Silicon Valley și Washington, încetinind ritmul modernizării. Puținele agenții guvernamentale care au ca obiect încurajarea inovației- United States Digital Service, Futures Command a Armatei Statelor Unite, comisia Defense Innovation Board și inițiativele Defense Innovation Unit Experimental (DIUx) – sunt fragile și supuse riscului de reducere a personalului și finanțării, în funcție de cei care se perindă la putere. Washingtonul

consideră că are o relație tranzacțională cu Mafia G. Nici parlamentarii și nici Casa Albă nu încearcă cu adevărat să dezvolte relațiile necesare cu conducerea Mafiei G pentru a forma o coaliție pe termen lung în domeniul inteligenței artificiale. Mafia G, armata și guvernul Statelor Unite își dau târcoale fără să-și unească eforturile în interesul național.

Am dat voie ca egoul și obiceiurile să împiedice degajarea unui consens național în privința Chinei. Politicienii, reprezentații comerciali, jurnaliștii, informaticienii și profesorii au dezbătut problema relațiilor dintre China, Statele Unite și inteligența artificială în continuu, fiecare parte susținându-și convingerile preferate, tradiționale, fără să lase loc unor realități alternative. Suspecții de serviciu susțin că Xi nu va rămâne prea mult timp la putere, chiar și dacă s-au abrogat limitele privind numărul de mandate. Odată ce Xi va pierde puterea, toate planurile pe termen lung ale Chinei se vor nărui. Alții combat această perspectivă: Xi va uni poporul și partidul. Indiferent dacă moare înainte de vreme sau își cedează poziția unui succesor, Partidul Comunist Chinez va fi mai puternic și va duce la îndeplinirea planurile privind inteligența artificială. Discuția se poartă tot așa, înainte și înapoi: politicile industriale chineze nu vor avea nici un impact-sau vor provoca distrugerea economiei Statelor Unite. Armata Chinei amenință lumea occidentală-sau e doar un articol exagerat de care ne vom plictisi curând. Ar trebui să investim timp și bani într-o strategie națională privind inteligența artificială, știind că planurile Chinei ar putea de greș-sau mai bine ne-am păstra timpul și banii și am adopta o abordare bazată pe așteptare. Un singur aspect este acceptat de toată lumea: dacă America are cu adevărat probleme, Mafia G va fi obligată să ne ajute.

Politicienii și grupurile noastre de gândire susțin aceleași argumente obosite, fără să acționeze, menținându-se într-o stare de inacțiune. Aleg această variantă pentru că în America e greu să scapi de forța centripetă a profitului fără o intervenție puternică.

Am auzit povestea acestei inacțiuni de multe ori până acum. Am apărat statutul fumatului, dezbătând datele concrete despre cancer și continuând să vindem tutunul ca un accesoriu al femeilor la modă, un stimulent pentru muncitori și un remediu medicinal pentru oameni bolnavi. Nu am luat măsuri în privința încălzirii globale, certându-ne constant în jurul unor definiții. Dacă încălzirea globală e reală, cum de poate fi atât de *rece?* Ne-am resemnat să dezbatem etapele. Afirmațiile alarmante ale anilor '70 au devenit sumbre în anii '90 apoi apocaliptice în anii 2010, dar suntem cu toții în continuare aici. Cine zice că lucrurile vor deveni atât de rele în viitor ?

Schimbarea sistemică are un efect de compunere și se acumulează în decenii, nu în zile. Când ne vom da seama ca inacțiunea a reprezentat alegerea greșită, va fi fost deja prea târziu.

2029: Prizonieri digitali

În ultimul deceniu, ați fost stimulați să cumpărați tot soiul de tehnologii și sisteme bazate pe tehnologia artificială. Toate aparatele electrocasnice sunt dotate cu astfel de sisteme. Frigiderul ține cont de mâncarea pe care o răcește. Mașinile de spălat -chiar și

cele de la spălătorii- urmăresc ciclul de spălare și vă anunță când s-a terminat. Cuptorul se închide singur înainte ca curcanul să se usuce și să ia foc.

Există însă și o problemă pe care nu ați anticipat-o: nu aveți permisiunea de a bloca funcționarea unui sistem de inteligență artificială care se voia a fi „de ajutor". După ce puneți prăjiturile, berea, brânza și carnea în frigider- toate cumpărate pentru ziua meciului Super Bowl- fapta este consemnată în DDP. Numărul de porții și calorii depășește cel normal pentru numărul de oameni din casă, prin urmare inteligența artificială concluzionează că plănuiți să mâncați peste măsură. Ar putea fi dincolo de miezul nopții și poate că ați planificat deja să puneți mâine rufele la spălat și uscat înainte de a pleca la muncă, însă inteligența artificială a mașinii de spălat nu ia în calcul dorința voastră de a dormi. Ea sună o alarmă și vă trimite constant notificări, fără să se oprească, pentru a vă spune că a venit vremea să puneți hainele la uscat. Ați vrea să faceți un mezel de curcan, dar cuptorul nu vă permite, pentru că inteligența artificială a fost programată să permită doar o carne suculentă. (Sau, dacă vă puteți permite, puteți plăti în plus pentru funcționalitatea care să vă permită să faceți și mezeluri.)

Unele persoane au probleme legate de inteligența artificială, mai ales aparatura de bucătărie, și de regulă dimineața. Panourile de control se sting, intermitent, ceea ce din nefericire are ca efect închiderea ușii, împiedicându-vă să puteți mânca. Mașina de spălat vase se oprește subit, la mijlocul ciclului, ținând paharele și argintăria să bălteasca într-o apă soioasă, plină de grăsimi. Volumul difuzorului inteligent devine dintr-o dată foarte ridicat, făcând imposibilă conversația cu membrii familiei la micul dejun. Și voi, și

zeci de mii de alți consumatori reclamați aceste defecțiuni, iar Mafia G însărcinează de fiecare dată câțiva manageri de produs să studieze problema. Jurnaliștii atribuie aceste defecte „felului dubios" în care „inteligența artificială se comportă uneori ciudat"

La început, atacurile par aleatorii, discrete. Vom da deci cu toții vina pe Google, Apple și Amazon pentru produsele defecte și serviciile detestabile de relații cu clienții. Experții în securitate informatică sunt apoi uluiți să descopere că toate aceste defecțiuni sunt de fapt interconectate. E un nou tip de atac cibernetic pornit din China, și efectuat cu ajutorul învățării mașinilor. Chinezii i-au dat și un nume: 被困, sau „bei kun", care în traducere înseamnă „prizonier." Hackerii, susținuți de guvernul chinez, au crezut că e o idee bună să ne atace la micul dejun și să ne țină mâncarea, băutura și tacâmurile blocate în aparatele noastre bazate pe inteligența artificială. Ei au un scop unic, și complicat: să sădească neîncrederea în Mafia G.

Microsoft și IBM continuă să existe, însă joacă un rol minor în spațiul inteligenței artificiale. Microsoft, o companie care publica odinioară studii de mare valoare în domeniul vizualizării digitale sau procesării limbajului natural, nu a reușit niciodată să-și canalizeze eforturile interne pentru a putea concura în domeniul inteligenței artificiale. Acum, compania își reduce dimensiunile și oferă în principal asistență pentru sistemele sale învechite: ceea ce a mai rămas din serviciile cloud Azure, Sharepoint, Skype și Outlook. În timp ce IBM Watson a găsit parteneri și clienți, serviciul

cloud al IBM, care ocupase mult timp de la mare distanță locul al treilea după Amazon și Microsoft, s-a micșorat odată ce Google a început să ofere prețuri competitive atât agențiilor guvernamentale cât și marilor companii. Celelalte divizii: centrele de date, stocare și semiconductori- nu au putut lupta împotriva concurenților din Taiwan, deveniți între timp cei mai mari furnizori din lume. Pentru companiile din Taiwan, principiul „O singură Chină" al Partidului Comunist Chinez se traduce printr-un avantaj semnificativ de piață, chiar dacă Beijingul le restricționează libertățile individuale. Politica industrială a Chinei a reușit să împiedice IBM să vândă în multe zone de pe glob.

Ce s-a ales de Facebook? După ani în care a promis că-și va spori securitatea și va oferi o transparență mai mare asupra felului în care partajează datele noastre, majoritatea utilizatorilor au ales alte platforme. Copiii din generația Alpha (copiii generației Y) au umplut rețeaua cu pozele lor, fără însă să-și facă conturi. Facebook o ia încet pe calea dispariției, după MySpace.

În 2035 interoperabilitatea reprezintă încă un punct slab în ecosistemul inteligenței artificiale din Occident, utilizatorii fiind forțați să accepte un sistem bazat pe o segregare de facto. Aparatele noastre sunt conectate la Google, Apple sau Amazon, așa că avem tendința de a cumpăra doar produsele și serviciile oferite de una dintre aceste trei companii- companii care ne-au vândut și toate aparatele bazate pe inteligență artificială din casele noastre- avem familii Google, familii Apple sau familii Amazon. Aceste etichete vin la pachet și cu prejudecăți nedorite.

Gospodăriile Apple au tendința de a fi mai în vârstă, și mai bogate. Acestea își pot permite toate aparatele Apple, disponibile

într-una din cele trei culori: un alb-argintiu, gri, sau onix închis. Ochelarii, toaletele şi frigiderele Apple continuă tradiţia lungă a produselor scumpe pe care oricine le poate folosi imediat după ce le scoate din cutie. DDP-urile Apple sunt dotate cu interfeţe vorbitoare şi două voci diferite, Joost (cu o „tonalitate unisex mai ridicată") sau Deva (cu o „tonalitate unisex mai gravă„). Uşurinţa de utilizare vine la pachet cu un cost. Inteligenţa artificială a Apple nu poate fi suprascrisă. Într-o casă Apple în care funcţionează aerul condiţionat, nu puteţi lăsa uşa deschisă mai mult de un minut fără ca sistemul să nu piuie fără încetare. Dacă senzorii din becuri detectează o cantitate suficientă de lumină de zi, sistemul Apple va împiedica acţionarea comutatorului de lumină.

Am văzut în avanpremieră sistemul Google pentru case conectate acum câteva decenii, la festivalul South By Southwest din 2018 din Austin, Texas. În vremea aceea, replica era „Make Google do it", iar modelele atrăgătoare invitau grupuri mici de oameni într-o casă cu trei etaje în care puteau să interacţioneze cu aparate cu ecrane bazate pe inteligenţa artificială şi blendere conectate. Sistemul Google e mai puţin intuitiv, dar foloseşte DDP-urile noastre mai bine-oferind niveluri diferite de deservire şi acces. Pentru cei care-şi pot permite tarifele de acces la o categorie mai bună şi au cunoştinţe mai bune în domeniul tehnologiei, Google Green oferă familiei capacitatea de a-şi debloca manual sistemele, putând conecta mai multe aparate – cum ar fi cafetiere, imprimante 3D, sisteme de irigaţii din grădină- la propria casă. Familiile Green au opţiunea de a renunţa la marketing şi reclame, deşi datele lor sunt în continuare strânse şi trimise unor terţi. Google Blue reprezintă o variantă accesibilă în care drepturile de deblocare sunt

limitate, însă familiile Blue primesc în continuare conținut de marketing. Google Yellow reprezintă nivelul cel mai de jos. Serviciul este gratuit, dar nu oferă capacitatea de schimbare a setărilor, oferă o selecție mică de aparate și o protecție limitată a datelor.

Amazon a ales o direcție interesantă dar care s-a dovedit în cele din urmă profitabilă. Câteva anunțuri făcute de Amazon în toamna lui 2018 au trecut în mare parte neobservate, cum ar fi lansarea cuptorului cu microunde Amazon Basics, care include o interfață vocală. Utilizatorii ar putea pune o pungă cu popcorn în microunde și să ceară Alexei să o prepare. Jurnaliștii din domeniul tehnologiei au considerat că acest cuptor cu microunde constituie un mod nou și prostesc de a o folosi pe Alexa, ignorând imaginea de ansamblu: sistemul fusese de fapt creat pentru a ne face dependenți de abonamentul la popcorn. Acel cuptor monitorizează atât ce încălzim dar și ce comandăm de pe platforma Amazon. O nouă cutie soseşte înainte ca cea veche să se termine.

Fiindcă Amazon a avut cea mai inteligentă abordare de colaborare cu agențiile federale, statale și locale- oferindu-le reduceri semnificative pe Amazon.com, acceptând cu răbdare condițiile de achiziții, construind și oferind mentenanța pentru servicii de cloud-, compania a reușit să devină platforma preferată pentru oferirea anumitor servicii sociale în Statele Unite. Amazon a descoperit astfel cum să extragă bani de la guvern.

Familiile cu venituri scăzute trăiesc în Amazon Housing, care a înlocuit programele de locuințe ale autorităților din Statele Unite. După orice standard, locuințele sunt mult mai bune decât cele oferite anterior prin programele guvernamentale. Casele Amazon sunt dotate în fiecare cameră cu aparate conectate. Fostul

program național de asistență nutrițională (cunoscut sub numele de Programul Bonurilor de Masă) este oferit în prezent prin Amazon, care oferă produse electrocasnice marca Amazon la prețuri reduse semnificativ, precum și alimente, cosmetice și cărți. Nu e de mirare că acest program funcționează fără niciun fel de problemă. Nu există niciodată întârzieri în distribuția fondurilor, contul poate fi accesat cu ușurință și toate tranzacțiile pot fi efectuate fără vreo coadă la un ghișeu. Cei care locuiesc în casele Amazon Homes trebuie să-și cumpere majoritatea produselor prin Amazon, în timp ce datele pe care le produc sunt filtrate și monetizate în diferite scopuri. Inteligența artificială a Amazon este cea mai răspândită, urmărind peste tot familiile Amazon pentru a colecta date comportamentale valoroase.

Imposibilitatea interconectării diferitelor structuri și sisteme de inteligență artificială a condus la o segregare în funcție de DDP și locuință, motiv pentru care avem în prezent un sistem de caste digitale. Alegând Google, Apple sau Amazon, suntem forțați să aliniem valorile familiei cu cele ale corporației. Familiile Apple sunt bogate, poate cu cunoștințe mai modeste în domeniul inteligenței artificiale și trăiesc în case frumoase. Familiile Google ar putea fi bogate și cu o înțelegere mai mare asupra tehnologiei, sau poate din clasa de mijloc și cu o toleranță mai mare față de marketing, sau acceptă faptul că nu e atât de important ca în viață să ai la dispoziție multe alegeri. Despre familiile Amazon se poate spune pur și simplu că sunt sărace, chiar dacă au acces gratuit la aparate cu puteri sporite.

Familiile sunt blocate în funcție de DDP-uri, iar etichetarea le urmărește. E mai ușor pentru o familie Google Yellow să treacă la nivelul Blue sau Green decât pentru un membru dintr-o familie Amazon să se conecteze la sistemul Apple. Acesta este motivul pentru care majoritatea familiilor au ales Google, atunci când au avut ocazia. Statutul este vizibil pentru orice inteligență artificială cu care interacționați. Taxiurile care se conduc singure, oferite de Lyft, uber și CitiCar, nu preiau călători Amazon prea ușor, iar mașinile alocate lor au tendința de a nu fi la fel de bune. Mașinile Waymo deservesc doar clienți Google. Pentru cei din categoria Green, mașina oferă temperatura dorită de călător și iluminatul în culorile preferate, conducând de-a lungul traseelor preferate de călător. Cei din categoria Yellow trebuie să îndure reclame pe toată durata călătoriei.

Reclamele nu sunt singura durere de cap pentru clienții Google din categoria Yellow. Un dezavantaj al tuturor aparatelor subvenționate (sau gratuite) oferite familiilor Google Blue, Yellow și Amazon este că alertele inteligenței artificiale legate de sănătate și bunăstare nu pot fi deconectate, iar aceste sisteme monitorizează, diagnostichează și sugestionează în mod constant. Când au fost inițial construite, informaticienii au trebuit să definească din necesitate sănătatea și bunăstarea pe baza unor criterii rigide. Acum, valorile triburilor inteligenței artificiale din trecut sunt o amintire oprimantă dintr-o epocă mai simplă. Nerespectarea alertelor legate de sănătate și bunăstare atrage multiple consecințe.

Mai țineți minte sistemul Amazon Locker, acel dulap pe care le foloseați în urmă cu mulți ani pentru a vă ridica cumpărăturile

din aplicația Amazon și de pe site-ul Amazon? El fac parte acum din casele Amazon. Departamentul de Sănătate și Servicii Sociale al Statelor Unite a considerat că sugestionarea reprezenta o variantă inteligentă de a îmbunătăți sănătatea și bunăstarea oamenilor săraci, așa că acest departament a elaborat noi politici prin care solicită tuturor ocupanților din locuințele sociale să fie dotați cu tehnologia Locker. Aceste dulapuri ar putea părea ca niște cămări, frigidere sau șifoniere obișnuite, însă ele funcționează ca niște jurii bazate pe inteligența artificială. Dacă un client Amazon Housing nu și-a îndeplinit norma de exerciții fizice, sistemul Locker va decide să-i încuie congelatorul și nu-l va lăsa să mănânce înghețată.

Simțim consecințele negative ale lucrurilor care ne oferă plăcerea și în afara caselor noastre Apple, Amazon și Google. Bordelurile tehnologizate, în care lucrează roboți sexuali bazați pe inteligența artificială, au căpătat o acceptare socială deoarece oferă o alternativă igienică la raporturile sexuale cu alți oameni. Bordelurile operează pe propriile platforme și necesită un abonament care vă permite să construiți și educați o personalitate bazată pe inteligență artificială. (Sau chiar personalități, pentru cei care-și permit pachetul premium). Alegeți pur și simplu un corp și vă uitați în ochii lui- camerele inteligente de dimensiuni reduse vă recunosc fața. Odată ce robotul se trezește, el discută cu voi de parcă nu ar fi trecut timpul și reacționează la toate dorințele și ordinele voastre. Sexul obișnuit, cu persoane normale vă dezamăgește.

Nu e imposibil să vă căsătoriţi în afara castei-din când în când o persoană Amazon se poate căsători cu un membru al unei familii Apple-însă dictonul cu „persoanele diferite se atrag" nu se mai aplică. Toate serviciile electronice matrimoniale ne aleg parteneri în funcţie de DDP şi de statut. Pe de o parte, nu mai suferim de tirania alegerii, deoarece sistemele de selecţie bazate pe inteligenţa artificială au redus dramatic numărul potenţialilor parteneri. Cu toate acestea, anumite alegeri care ne confereau un caracter distinct omenesc-iubirile de-o vară, relaţiile cu persoane pe care părinţii nu le agreează- sunt mai greu de obţinut acum. În America, societatea începe să aibă un caracter deranjant, huxleyan, pe măsură ce avem relaţii, ne căsătorim şi facem copii cu persoane din aceeaşi castă, fie ea Apple, Google Blue sau Amazon.

După cum s-a anticipat, inteligenţa artificială şi automatizarea încep să elimine munca omenească- mult mai multe locuri de muncă decât ne-am aşteptat. Şomajul tehnologic care se contura de multă vreme la orizont a devenit o realitate, însă nu după cum ne-am imaginat cu toţii. Eram pregătiţi ca şoferii de camion sau muncitorii să rămână fără lucru, însă previziunile s-au dovedit greşite. Am tot presupus că roboţii vor elimina slujbele slab calificate, dar se pare că construirea unor roboţi capabili de munca fizică e o sarcină mult mai grea decât ne-am imaginat, în vreme ce munca cognitivă e mult mai uşor de programat şi produs. În mod ironic, muncitorii din domeniul cunoaşterii sunt cei de care nu mai e nevoie.

Ca urmare, America şi aliaţii săi au o nevoie critică, imediată de muncitori în domeniile pe care le-am crezut pe cale de dispariţie.

Pur şi simplu nu avem suficienţi instalatori, electricieni şi tâmplari. Roboţii nu oferă aceeaşi senzaţie de omenesc de care avem nevoie, motiv pentru care maseurii, manichiuristele, cosmeticienele sau bărbierii sunt la mare căutare. Împotriva automatizării se manifestă o reacţie negativă. Majoritatea oamenilor nu vor ca un robot să le pregătească băuturile sau cafeaua. Vrem să consumăm în preajma oamenilor. Faptul că am preferat disciplinele reale în favoarea celor umaniste a reprezentat o abordare întrucâtva greşită. Muncitorii sunt cei care mai au de lucru, nu informaticienii sau programatorii. Tocilarii au construit programe care i-au lăsat fără locurile de muncă.

Fără să vrea, Google, Amazon şi Apple au creat un trident în cadrul inteligenţei artificiale, puterea fiind consolidată masiv. În America şi în ţările aliate, avem la dispoziţie produse spectaculoase, dar foarte puţine alegeri. Spre exemplu, putem plăti pentru a avea ochelari OmniVision cu vedere sporită, dincolo de limitele vederii omeneşti. Doar două companii produc ochelarii: Google şi Apple. Dacă nu vă plac cum arată, sau dacă ei nu se potrivesc cu forma unică a feţei sau a urechilor, nu aveţi noroc. Amazon vinde orice vă puteţi imagina, dar produsele esenţiale zilnice reprezintă mărcile proprii ale companiei. În ţările democratice din toată lumea, există o abundenţă de mărfuri, dar varietatea şi alegerea de pe piaţă sunt strâns controlate. Deşi avem bani de cheltuială, avem o putere de cumpărare foarte redusă. Într-un mod bizar, ne aducem aminte de fosta Uniune Sovietică.

Salesforce, compania care oferă programe de relaţii cu clienţii şi cloud computing a încheiat un parteneriat cu Google, Amazon

și Apple acum mult timp pentru a construi un modul educațional pentru DDP-uri. Acum, testarea riguroasă și clasificarea care erau emblematice pentru educația din America în anii 80 și 90 au devenit din nou la modă. Capacitățile noastre cognitive sunt evaluate înainte de grădiniță, iar reușitele și progresele sunt monitorizate pe tot parcursul vieții.

Metricii și optimizarea au reprezentat dintotdeauna valorile cheie ale Salesforce, iar acum ele reprezintă valorile cheie ale educației în America. Îngrijorați că am înlocuit înțelepciunea cu acumularea unor informații devenite între timp inutile, conducătorii din domeniul educației au abandonat programa comună de bază în favoarea unui nou concept. Dată fiind criza forței de muncă din america, elevii sunt împărțiți în cadrul examenelor de admitere la grădiniță în două categorii: filiera vocațională sau cea de conducere. Elevii din filiera vocațională sunt educați să fie agili în mai multe discipline, în timp ce elevii din filiera de conducere primesc cursuri de gândire critică și management. Aptitudinile managementului de mijloc nu mai sunt necesare deoarece în mare parte activitatea managerilor de mijloc și a lucrătorilor din domeniul cunoașterii este îndeplinită acum de inteligența artificială.

Dat fiind șomajul apărut în domenii unde nu ne așteptam, și infracționalitatea a crescut-însă nu din motivele la care v-ați fi așteptat. Programele de ordine publică bazate pe inteligență artificială nu au funcționat cum trebuie, așa că statisticile privind criminalitatea nu reflectă în mod precis lumea reală. Algoritmii construiți de triburile inteligenței artificiale și instruiți pe baza unui set limitat de date nu au învățat niciodată cum să identifice și să clasifice o persoană cu tulburări de gen- o persoană care nu

se identifică nici ca bărbat, nici ca femeie și care poate avea un aspect androgin sau o barbă dar și gene false. Ca urmare, sute de oameni care nu îndeplinesc criteriile unui anume gen sunt acuzați în mod eronat în fiecare zi că fură identitatea altora: când încearcă să plătească prin recunoaștere facială, când se plimbă prin birou sau când încearcă să efectueze un apel video. Pentru moment, singura soluție este să încerce să se conformeze unui anumit gen în timpul anumitor tranzacții. Ei/ele trebuie să-și pună o anumită perucă specifică genului, sau să de demachieze pentru a deveni temporar un el/ea în ochii inteligenței artificiale care procesează datele vizuale. E un memento umilitor și public al faptului că diversitatea nu a contat niciodată suficient de mult încât să reparăm un sistem defect.

Inteligența artificială conferă Google, Apple și Amazon o putere economică uriașă- iar Chinei o putere militară și geopolitică de neimaginat. La finalul anilor 2030, ne-am dat seama că inteligența artificială s-a dezvoltat de-a lungul unor traiectorii paralele, sprijinind capitalismul în Occident, și varianta chineză a comunismului în Asia, Africa și America latină. America și aliații săi, care sărbătoreau odinioară reușitele Mafiei-G, trăiesc acum într-un sistem totalitar bazat pe inteligență artificială. Cetățenii chinezi și ai țărilor care beneficiază de investițiile directe și infrastructura chineză constată că și ei trăiesc sub un aparat omniprezent de pedepse și recompense bazat pe inteligența artificială.

2049: Granițe biometrice și avorturi cu nano-boți

Mafia-G a devenit grupul GAA: Google, Apple și Amazon. Facebook a fost prima dintre companii care și-a declarat falimentul. Rămășițele Microsoft și IBM au fost achiziționate de Google.

Sărbătorim 100 de ani de la Revoluția Comunistă din China și proclamarea Republicii Populare Chineze (RPC) de către Mao Zedung. Ample festivități sunt anunțate în toate țările satelit ale Chinei pentru onora memoria defunctului Xi Jinping și a sărbători ascensiunea dinastiei Réngōng Zhìnéng (Inteligența Artificială).

Toată omenirea este acum înconjurată de sisteme de inteligență artificială, care ar fi trebuit să ne ajute să ducem vieți mai libere și mai fericite. Încă de la început, triburile inteligenței artificiale din Statele Unite au spus că vor să trăim cât mai bine, să avem activități creative și să colaborăm în privința celor mai mari provocări ale omenirii. A fost un ideal utopic născut în mediul izolat din Silicon Valley, iar creatorii acestui ideal pierduseră complet orice legătură cu lumea din afară.

Toate aceste sisteme au fost create să ne facă viața mai ușoară dar au reușit în schimb să ne facă mai leneși. Ele ne-au subminat senzația de productivitate, de a avea un scop. Ne bazăm pe aceste sisteme pentru a decide în locul nostru. Ne resemnăm la niște alegeri limitate. Trecem prin etapele pre-programate ale traiului cotidian, optimizate de inteligența artificială generală pentru toți locuitorii planetei.

Multe sisteme IAG au evoluat pentru a concura, nu pentru a colabora. Atacurile Chinei asupra interfețelor aparatelor de acum două decenii ne par atât de simpliste, de benigne. Trăiți într-o în-

chisoare bazată pe inteligența artificială pe care v-ați construit-o singuri. Pierdeți controlul constant asupra cuptorului, șifonierului sau băii, și nici nu vă mai deranjați să vă mai luptați. Nu mai are sens. Reacția normală, vi s-a spus, este să așteptați ca problema să se rezolve. Casele Google Green și Apple pot achiziționa un abonament suplimentar de top, care permite teoretic accesul la o IAG de reparație care să suprascrie codul greșit-însă IAG-urile sunt blocate într-o buclă a auto-perfecționării. Toți banii din lume nu pot repara o casă afectată de o defecțiune sistemică.

Concentrarea averilor a permis GAA să obțină rezultate uimitoare în domeniul sănătății. Google a fost prima companie care a lansat un proiect pilot cu roboți microscopici, injectabili, capabili să livreze medicamentele doar unei anume părți a corpului sau să asiste în micro-chirurgie. Nanoboții sunt fabricați acum sub multe forme diferite. Spre exemplu, există un robot autonom, molecular, construit dintr-o singură fâșie de ADN care tratează interiorul corpului omenesc ca pe un depozit de distribuție. Acest nanobot se poate plimba, luând molecule dintr-un loc și depunându-le în alte locuri stabilite. Un alt tip de nanobot, care se deplasează pe baza unor bule de aer, poate să administreze cantități microscopice de medicamente fără să provoace rău. Apariția acestor nanoboți pe piață, care partajează informații cu DDP-urile noastre, a înlocuit terapiile și medicațiile generalizate, tratând suferințele noastre distincte fără să provoace efecte secundare.

Acum că atât Amazon cât și Apple oferă servicii medicale personalizate, majoritatea oamenilor s-au injectat de bună voie cu

nano-boți organici. Chiar și familiile Amazon au acces printr-un program subvenționat aprobat de guvernul Statelor Unite. Nanoboții continuă să ne monitorizeze și să ne trateze, astfel încât speranța de viață a unui american de rând a crescut de la 76,1 ani în 2019 la 99,7 ani în zilele noastre.[184]

Nu a durat prea mult până când am descoperit și potențialele efecte negative ale IAG injectabile. Nanoboții au făcut exact ceea ce au dorit creatorii lor. Ei s-au comportat imprevizibil și au început să învețe. Dacă ne gândim în urmă, construirea și instruirea sistemelor de inteligență artificială pentru a face lucruri la care nu ne-am fi gândit niciodată a fost un obiectiv principal al triburilor inteligenței artificiale. Aceasta a fost cheia rezolvării problemelor spinoase pe care omenirea nu le-ar fi putut confrunta de una singură. Când AlphaGo Zero a luat decizii strategice autonome în urmă cu decenii, am considerat că inteligența artificială atingea noi culmi. În interiorul corpurilor noastre, însă, nanoboții și IAG-urile care le controlează se îmbunătățesc singure și au mai multă putere de decizie decât ne-am fi dorit.

Avem acum o nouă *himeră economică* a oamenilor. Casele Apple și Google Green pot accesa super-puteri, obținând o cogniție superioară, un miros extrasenzorial și o atingere îmbunătățită.

Casele Google Blue, Yellow și Amazon doar că nu au acces la aceste îmbunătățiri ci au și anumite restricții biologice. Când o persoană rămâne însărcinată, IAG-urile rulează constant modele predictive pentru a stabili sănătatea și viabilitatea fetusului. Ceea ce nimeni nu anticipase era ca IAG-urile să-și ducă obiectivele la extrem. Din cauză că obiectivul programat era de a sprijini oame-

nii să aducă pe lume fetuşi sănătoşi, IAG-urile au început să caute anomalii ale ţesutului la făt. Dacă se detecta o astfel de anomalie, IAG avorta automat fătul, fără să ofere părintelui opţiunea de a alege dacă păstrează copilul.

În mod similar, nanoboţii monitorizează procesul de îmbătrânire, calculând când devine continuarea vieţii mai dureroasă decât moartea. Odată ce începeţi să aveţi nevoie de îngrijire la domiciliu şi deveniţi o povară asupra resurselor statului, IAG-urile intervin, inducând o moarte confortabilă, pentru ca nici voi, şi nici rudele voastră, să nu trebuiască să decidă când e vremea să muriţi.

Legile ţărilor GAA au fost depăşite o dată ce IAG-urile s-au îmbunătăţit şi au putut crea mecanismele care stabilesc cine trăieşte şi cine moare. Guvernele s-au grăbit să creeze un cadru juridic şi reglementări care nu ajută însă la nimic. Interzicerea nanoboţilor ar presupune revenirea la medicina tradiţională, iar marile companii farmaceutice care fabricau medicamentele de care am fi avut nevoie au dispărut. Chiar şi cele mai optimiste proiecţii ne arată că repornirea fostelor sisteme sanitare ar presupune un efort de un deceniu sau chiar mai bine- iar, între timp, milioane de oameni ar suferi de multiple boli.

Cercetătorii au dezvoltat în schimb un nou tip de nanobot bazat pe IAG care poate controla alţi nanoboţi din interiorul corpului-imitând felul în care leucocitele luptă cu un virus. La fel ca inteligenţa artificială, ideea a fost inspirată de biologia omului. Pe măsură ce corpurile noastre se luptă cu nanoboţii IAG indezirabili, simptomele sunt mult mai neplăcute decât cele ale unei gripe, şi mult mai periculoase.

Marile companii sunt conduse acum de directori generali reprezentați de inteligența artificială- sunt cei care calculează riscurile și oportunitățile strategice. Directorii umani lucrează alături de aceste entități, reprezentând „fața" companiei. Companiile de dimensiuni mici și medii- restaurantele, service-urile auto, saloanele cosmetice- sunt toate partenere ale uneia din grupul GAA. Pe lângă DDP-urile personale și ale caselor, fiecare firmă și ONG este acum înregistrată într-un Dosar de Date ale Organizației.

Cu toate acestea, foarte mulți oameni din America și din țările aliate nu au un loc de muncă. În absența unei plase sociale de siguranță, economiile occidentale se află într-un declin prelungit, deoarece nu s-au putut redresa după valurile neanticipate de șomaj tehnologic. Astfel s-au creat vulnerabilități- și o fereastră pentru investițiile chineze. În curând, politicienii sunt forțați să aleagă între idealurile democratice și viabilitatea economică- o decizie deosebit de grea pentru persoane care se confruntă cu alegerile și presiunea de a rezolva problemele domestice urgente.

Ca răspuns, Statele Unite încearcă să limiteze expansiunea Chinei prin blocade comerciale, sancțiuni secundare și alte tactici diplomatice. America constată însă că nu mai are influența geopolitică de odinioară. Liderii americani au petrecut prea mult timp deliberând în loc să acționeze în privința Chinei. Ei au călătorit prea puțin în America Latină, Africa și Asia de Sud-Est. Ei nu și-au câștigat niciodată încrederea, favoarea și prietenia omologilor lor din străinătate.

Demersurile chineze în domeniul inteligenței artificiale devin din ce în ce mai importante. Scorul de armonie socială este folosit

în peste 100 de țări, înlocuind documentele tradiționale de călătorie. China excelează în construcția zidurilor, iar Marele Zid al Inteligenței Artificiale nu face excepție de la această regulă. Zidul oferă atât o barieră de protecție împotriva străinilor, dar și o metodă prin care datele tuturor pot fi extrase și analizate. Cei care au un scor al armoniei sociale suficient de ridicat primesc acces nelimitat (dar monitorizat, desigur) prin Marele Zid al Inteligenței Artificiale către oricare dintre țările conectate la rețeaua Chinei. China a înființat granițe biometrice cu recunoaștere facială pentru a putea stabili cine poate intra sau ieși. Nu mai există poliție de frontieră, și nici pașapoarte care să fie ștampilate.

La granița de sud a Statelor Unite există acum un zid. El este un zid de senzori și a fost construit pe pământ mexican de către chinezi, pentru a ne izola. Deoarece americanii nu au acces la sistemul de credite sociale, vi se refuză intrarea în fostele destinații preferate de vacanță: Bahamas, Jamaica, Cancun, Playa del Carmen, Cozumel, Costa Rica și Aruba. Dacă încercați să treceți ilegal o frontieră biometrică, o IAG emite un atac sonic care provoacă greață, contuzii, sângerări auriculare și stres psihologic.

Americanii și aliații noștri sunt izolați de prietenii și rudele din țările satelit ale Chinei, deoarece PCC controlează toată infrastructura rețelelor din aceste țări. Dacă aveți nevoie să căutați pe cineva dintr-o țară aflată sub controlul PCC, trebuie să treceți prin China ca intermediar, știind că toate cuvintele sunt ascultate.

Grupul GAA a format în cele din urmă o coaliție cu guvernul Statelor Unite și aliații rămași. Date fiind restricțiile economice și de călătorie impuse de China, există puține resurse financiare prin care să poată fi găsită o soluție acceptabilă. Se ia decizia de a dez-

volta o IAG care să rezolve problema Chinei pentru noi. Sistemul vede însă doar două soluții: să cedăm Chinei, sau să reducem gradual populația globului.

2069: Anihilarea digitală

În timp ce China se preocupa de planificarea pe termen lung și o strategie națională în domeniul inteligenței artificiale, Statele Unite erau preocupate de aparate și dolari.

China nu mai are acum nevoie de Statele Unite ca partener comercial și nu are nevoie de drepturile noastre de autor. China a construit o rețea de peste 150 de țări care operează pe baza principiilor directoare ale politicii globale chineze. La schimb pentru obediența lor, aceste țări primesc acces la rețea, capacitatea de a face comerț și un sistem financiar stabil sprijinit de Beijing. Cetățenii lor pot călători liber în spațiul chinez, cu condiția să aibă un scor social suficient de ridicat.

Libertatea de mișcare-un drept pe care americanii considerau că li se cuvine- nu ne-a lipsit niciodată mai mult. Motivul e că America trece, la fel ca multe alte țări, printr-o criză a populației. Pe planeta noastră trăiesc peste 10 miliarde de oameni. Se nasc prea mulți oameni, și prea repede, și cu toate acestea insistăm să ne prelungim viața până dincolo de 120 de ani.

Populația planetei constituie o problemă pentru că nu am combătut suficient de repede încălzirea globală, nici măcar după ce China a preluat răspunderea sustenabilității și a protecției mediului. Două treimi din terenurile arabile au fost pierdute. Deși facem eforturi ca să construim ferme subterane în America, nu

putem creşte mâncarea suficient de repede pentru a hrăni populaţia. Sancţiunile globale au blocat rutele comerciale şi ne-au izolat de naţiunile care exportă mâncare, dar şi China şi sateliţii săi au dificultăţi.

Într-o bună zi, familiile Apple suferă de o boală misterioasă. DDP-urile lor prezintă o anomalie, fără să ofere amănunte sau detalii concrete. La început, credem că această ultima versiune de nanoboţi e defectă, aşa că managerii de produs se grăbesc să ofere variante de IAG care să repare deficienţele. Boala loveşte apoi casele Google- nu doar în America, ci şi în toate locuinţele din afara spaţiului chinez. Boala misterioasă se agravează rapid.

China a construit o SIA care are un singur scop: exterminarea populaţiilor Americii şi a aliaţilor acesteia. Ţările din arealul chinezesc au nevoie de resursele rămase ale planete, iar Beijingul a calculat că unica metodă de a supravieţui este de a ne lua aceste resurse.

Ceea ce urmează este mai grav decât orice bombă inventată vreodată. Bombele au efecte imediat şi distrugător. Anihilarea prin inteligenţa artificială este înceată şi ireversibilă. Rămâneţi neajutoraţi în timp ce corpurile copiilor vă mor în braţe. Vedeţi cum colegii se prăbuşesc la birou. Simţiţi o durere ascuţită. Vă ia ameţeala. Respiraţi o ultimă respiraţie, repede, superficial.

Iată sfârşitul Americii.

Iată sfârşitul aliaţilor Americii.

Iată sfârşitul democraţiei.

Dinastia Réngōng Zhìnéng preia puterea, brutal, ireversibil şi absolut.

Există în prezent semnale care validează fiecare dintre aceste trei scenarii. Trebuie să alegem o variantă acum. Voi trebuie să alegeți. Vă cer să alegeți scenariul optimist, și să construiți un viitor mai bun pentru inteligența artificială, și pentru omenire.

PARTEA A III-A:
Rezolvarea problemelor

CAPITOLUL OPT

PIETRICELE ȘI BOLOVANI: CUM SĂ REPARĂM VIITORUL INTELIGENȚEI ARTIFICIALE

Concluziile ultimului capitol ar putea părea extreme și improbabile. Există însă deja semnale care ne spun că dacă nu alegem un viitor în care cei Nouă Titani să fie stimulați să colaboreze în interesul omenirii, este foarte probabil să ajungem să trăim într-o lume care să semene cu dinastia Réngōng Zhìnéng.

Cred că scenariul optimist- sau o variantă apropiată de acesta- e la îndemână. Este posibil ca inteligența artificială să-și îndeplinească scopul superior și potențialul, spre beneficiul triburilor inteligenței artificiale dar și al tuturor oamenilor. Pe măsură ce evoluează, inteligența artificială poate deservi atât oamenii din China și din Statele Unite, cât și toți aliații noștri. Ea ne poate ajuta să ducem o viață mai sănătoasă, să reducem ecarturile economice și să ne sporească siguranța, pe stradă sau în casă. Inteligența artificială ne poate oferi puterea de a descoperi tainele întrebărilor celor mai profunde ale omenirii, cum ar fi originile și modul în care a apărut viața. În cadrul acestui proces, inteligența artificială

ne va putea ului și distra, creând lumi virtuale pe care nu ni le-am fi putut niciodată imagina, scriind cântece care să ne inspire și creând noi experiențe distractive și care să ne împlinească. Nu vom putea avea însă niciunul dintre aceste beneficii fără planificare, fără un angajament de a pune osul la muncă și fără un exemplu curajos al tuturor celor implicați în grupurile care dezvoltă inteligența artificială.

Tehnologiile sigure și benefice nu sunt rezultatul speranțelor și al hazardului. Ele sunt produse ca urmare a contribuțiilor unor lideri curajoși, și a unor colaborări dedicate, de lungă durată. Cei Nouă Titani se află sub o presiune imensă -din partea sistemului financiar în Statele Unite, respectiv a Beijingului în China- de a îndeplini așteptări pe termen scurt, chiar dacă periclitează astfel viitorul nostru. Trebuie să emancipăm cei Nouă Titani și să le insuflăm curajul de a schimba traiectoria inteligenței artificiale, deoarece în absența unui sprijin deplin din partea noastră, aceste companii nu pot și nu vor dori să reușească pe cont propriu.

Vint Cerf, unul dintre cei care a elaborat primele protocoale și arhitectura internetului modern, folosește o parabolă pentru a explica de ce conducerea curajoasă este esențială în epoca tehnologiilor emergente, precum inteligența artificială. [185] Imaginați-vă că trăiți într-o comunitate minusculă într-o vale înconjurată de munți. În vârful unui munte îndepărtat se află un bolovan uriaș. A stat acolo de mult timp și nu s-a mișcat niciodată, așa că din punctul de vedere al comunității voastre, e doar parte din peisaj. Într-o bună zi, observați că bolovanul uriaș pare instabil-că se pregătește să se rostogolească la vale, cu o viteză și o putere din ce în ce mai mare, și că va distruge comunitatea voastră și pe toți

membrii ei. De fapt, vă dați seama că poate toată viața ați ignorat mișcarea lui. Acel gigantic bolovan s-a mișcat dintotdeauna, puțin câte puțin, însă nu ați acordat atenția voastră deplină schimbărilor subtile, minuscule, de fiecare zi: o mică schimbare a umbrei pe care o aruncă, distanța dintre el și următorul munte, sunetul aproape imperceptibil pe care-l produce în timp ce strivește pământul de sub el. Vă dați seama că o singură persoană nu poate urca pe munte și să oprească bolovanul gigantic de una singură. Sunteți prea mic, și bolovanul e prea mare.

Vă dați apoi seama că dacă puteți găsi o pietricică pe care să o plasați într-un loc bun, ea va încetini mișcarea bolovanului și îi va devia un pic traiectoria. O singură pietricică nu va împiedica bolovanul să distrugă satul, așa că cereți întregii comunități să vi se alăture. Cu pietricelele în mână, fiecare persoană urcă muntele și se pregătește pentru venirea bolovanului- există colaborare, și comunicare, și un plan de a confrunta bolovanul pe măsură ce coboară. Oamenii și pietricelele lor- nu un bolovan mai mare- reușesc să facă diferența.

Ceea ce urmează este o serie de pietricele. Voi începe foarte larg, subliniind motivele pentru constituirea unei comisii la nivel global care să supervizeze parcursul inteligenței artificiale și nevoia noastră imediată de a avea norme și standarde. Voi explica apoi ce schimbări vor trebui luate de către guvernele Chinei și Statelor Unite. După aceea, voi arunca o privire mai în amănunt asupra nevoii celor Nouă Titani de a-și reforma practicile. Mă voi concentra apoi doar asupra triburilor inteligenței artificiale și a universităților în care acestea se formează și voi explica în amănunt ce schimbări trebuie făcute chiar acum. În cele din urmă, voi

explica rolul pe care și voi îl puteți juca în modelarea viitorului inteligenței artificiale.

Viitorul în care cu toții ne vom *dori* să trăim nu se va împlini de la sine și complet. Trebuie să fim curajoși. Trebuie să ne asumăm răspunderea pentru acțiunile noastre.

Schimbarea sistemică globală: argumentul pentru înființarea GAIA

În scenariul optimist, un amestec divers de lideri din economiile cele mai avansate își unesc eforturile cu Mafia G pentru a înființa Alianța Globală de Augmentare a Inteligenței, sau GAIA. Acest corp internațional include cercetători din domeniul inteligenței artificiale, sociologi, economiști, teoreticieni ai jocurilor, futurologi, politicieni din toate statele membre. Membrii GAIA reflectă o diversitate socioeconomică, de gen, rasă, religie, politică și sexuală. Ei sunt de acord să faciliteze și să coopereze în inițiative comune și politici legate de inteligența artificială, reușind cu timpul să exercite suficientă influență și control încât să împiedice o apocalipsă-fie ea datorată IAG, SIA sau folosirii inteligenței artificiale în China pentru a oprima cetățenii.

Cea mai bună modalitate de a îndeplini schimbarea sistemică este de a vedea înființarea GAIA cât de curând posibil, iar aceasta ar trebui să aibă sediul pe un teren neutru, în apropierea unui centru de inteligență artificială. Cel mai bun loc în care GAIA ar trebui să-și aibă sediul este în Montreal, Canada. Mai întâi, Montreal găzduiește multe laboratoare de cercetare și cercetători în domeniul învățării profunde. Dacă presupunem că tranziția de la IAÎ la

IAG va include învățarea profundă și rețele neurale profunde, rezultă că GAIA ar trebui să își aibă sediul în apropiere de locul unde se desfășoară această activitate. În al doilea rând, în regimul prim-ministrului Justin Trudeau, guvernul canadian a alocat deja personalul și resursele financiare pentru a explora viitorul inteligenței artificiale. În anii 2017 și 2018, Trudeau nu s-a rezumat să vorbească despre inteligența artificială, ci a poziționat Canada pentru a putea contribui la modelarea reglementărilor și principiilor care stabilesc evoluția inteligenței artificiale. În al treilea rând, Canada reprezintă un teritoriu geopolitic neutru pentru inteligența artificială- este departe atât de Silicon Valley cât și de Beijing.

Poate părea imposibil ca guvernele lumii să se unească în jurul unei cauze, date fiind tensiunile politice și instabilitatea geopolitică din ultimii ani. Există însă și un precedent. La finalul celui de-al doilea război mondial, când tensiunile erau încă ridicate, sute de delegați din toate țările aliate s-au adunat la Bretton Woods, în statul New Hampshire, pentru a construi structurile financiare care au permis economiei globale să meargă mai departe. Această colaborare a fost centrată pe oameni-ea a rezultat într-un viitor în care oamenii și țările au putut reconstrui și să caute prosperitatea. Țările GAIA ar trebui să colaboreze în domeniul cadrelor de referință, standardelor și bunelor practici în domeniul inteligenței artificiale. Deși este improbabil ca China să se alăture, liderii PCC și ai grupului BAT ar trebui să fie invitați.

Înainte de orice, GAIA va trebui să stabilească o modalitate prin care să garanteze drepturile elementare ale omului în epoca inteligenței artificiale. Când vorbim despre inteligență artificială și etică, avem tendința să ne gândim la cele trei legi ale roboticii

ale lui Isaac Asimov, pe care le-a publicat într-o nuvelă din 1942 intitulată „Runaround."[186] Era o poveste despre un computer umanoid, nu despre inteligența artificială, însă cu toate acestea aceste legi ne-au inspirat gândirea etică atât de mulți ani mai târziu. După cum am amintit în Capitolul 1, regulile lui Asimov sunt sunt: (1) Un robot nu poate răni un om, sau prin inacțiune să permită ca un om să sufere. (2)Un robot trebuie să asculte ordinele pe care i le dau oamenii, mai puțin în cazurile în care aceste ordine s-ar afla în conflict cu Prima Lege.(3) Un robot trebuie să-și protejeze existența atât timp cât această protecție nu intră în conflict cu Prima sau a Doua Lege.

Când Asimov a publicat ulterior o colecție de nuvele într-o carte intitulată *I, Robot*, el a adăugat o lege „Zero", care să le supună pe toate celelalte: (0) roboții nu pot face rău oamenilor. Asimov a fost un scriitor talentat, clarvăzător-dar legile sale ale roboticii sunt prea generale pentru a servi drept principii directoare pentru viitorul inteligenței artificiale.

GAIA ar trebui în schimb să creeze un nou contract social între cetățeni și cei Nouă Titani (definiți în mod extins ca fiind Mafia G și grupul BAT precum și partenerii, investitorii și filialele acestora). Contractul trebuie să se bazeze pe încredere și colaborare. Membrii GAIA ar trebui să fie de acord că inteligența artificială trebuie să emancipeze cât mai mulți oameni din jurul lumii. Cei Nouă Titani ar trebui să acorde prioritate drepturilor omului și nu ar trebui să ne considere niște resurse care să fie exploatate pentru profit sau scopuri politice. Prosperitatea economică promisă de inteligența artificială și de cei Nouă

Titani ar trebui să sporească considerabil bunăstarea tuturor.

Rezultă așadar că Dosarele noastre de Date Personale ar trebui să poată fi portate între operatori și ar trebui să fie proprietatea noastră, nu a unor companii, sau consorții, sau națiuni. GAIA ar putea începe încă de azi să studieze cum se poate implementa această măsură, pentru că DDP-urile despre care este vorba în scenarii există deja în zilele noastre, în formă incipientă. Ele se numesc „Informații Personale Identificabile" sau IPI. Aceste IPI individuale sunt cele pe baza cărora funcționează aplicațiile din telefoanele noastre inteligente, rețelele de publicitate din site-uri, precum și recomandările care ne sugestionează pe ecrane. IPI sunt transmise unor sisteme folosite pentru a ne identifica și localiza. Felul în care sunt folosite depinde în întregime de capriciile companiilor și agențiilor guvernamentale care le accesează.

Înainte de a genera un nou contract social, GAIA trebuie să decidă cum pot fi folosite DDP-urile pentru a ajuta algoritmii de învățare ai mașinilor, și trebuie să definească ce anume constituie valorile de bază în epoca automatizării. Definirea clară a valorilor are o importanță esențială pentru că aceste valori sunt cele care vor fi în cele din urmă programate în datele de instruire, datele din lumea reală, sistemele de învățare și aplicațiile care alcătuiesc ecosistemul inteligenței artificiale.

Pentru a putea cataloga valorile noastre de bază, GAIA ar trebui să creeze un Atlas al Valorilor Omenirii, care să definească valorile noastre unice în funcție de diferite țări și culturi. Acest atlas nu ar trebui să fie unul static. Deoarece valorile se schimbă cu timpul, atlasul ar trebui să fie actualizat de țările membre. Putem căuta precedentul în biologie: Atlasul Celulelor Umane reprezintă o co-

laborare globală a comunității științifice, care include mii de experți din diferite domenii (genetică, inteligență artificială, programare, vizualizarea datelor, medicină, chimie și biologie). [187] Proiectul cataloghează fiecare tip de celulă din corpul uman, cartografiind tipologiile celulare cu locul unde se găsesc, urmărind istoria celulelor pe măsură ce evoluează și surprinzând caracteristicile celulelor pe durata vieții. Acest demers-costisitor, complicat, cronofag și continuu-va permite cercetătorilor să facă progrese semnificative, iar el este posibil doar grație unei colaborări substanțiale, la nivel mondial. Ar trebui să avem un atlas asemănător al valorilor umane, în care să fie implicați profesori, antropologi, sociologi, psihologi dar și oameni obișnuiți. Crearea Atlasului Valorilor Umane ar fi un demers scump, problematic și greoi- și ne-am lovi de multe contradicții, deoarece valorile unor culturi ar fi contrare cu ale altora. Cu toate acestea, în absența unui cadru și a unui set de standarde elementare, le cerem celor Nouă Titani și triburilor inteligenței artificiale să facă ceva imposibil-să ia în considerare toate perspectivele noastre și toate rezultatele posibile în ce privește grupuri distincte din cadrul societății, din fiecare țară din lume.

GAIA ar trebui să ia în considerare o structură de drepturi juridice care să pună în balanță libertățile individuale și binele comun. Ar fi de preferat ca structura dezvoltată să reprezinte bine idealurile dar să ofere o interpretare mai flexibilă, pe măsură ce inteligența artificială evoluează. Organizațiile membre ar trebui să poată demonstra că respectă această structură sau să fie amenințate cu excluderea din GAIA. Orice structură ar trebui să includă următoarele principii:

1. Omenirea ar trebui să se afle întotdeauna în centrul evoluției inteligenței artificiale.
2. Sistemele de inteligență artificială ar trebui să fie sigure. Ar trebui să putem verifica singuri cât de sigure sunt.
3. Cei Nouă Titani -inclusiv investitorii, angajații și guvernele pentru care lucrează- trebuie să optimizeze siguranța în favoarea vitezei. Orice echipă care lucrează la un sistem de inteligență artificială-chiar și cele din afara celor Nouă Titani- trebuie să acorde prioritate siguranței, și nu vitezei. Siguranța trebuie să poată fi demonstrată și confirmată de observatori independenți.
4. Dacă un sistem de inteligență artificială face rău, el trebuie să poată da socoteală privind erorile fiind necesar un proces decizional prin care pagubele să poată fi analizate și combătute.
5. Inteligența artificială trebuie să fie explicabilă. Sistemele ar trebui să dețină un soi de etichetă cu informații nutritive, prin care să explice ce fel de date de instruire au fost folosite, ce procese de învățare, datele din lumea reală folosite în aplicații și rezultatele anticipate. În ce privește sistemele sensibile sau protejate de drepturi de autor, transparența unei inteligențe artificiale ar trebui să poată fi evaluată și verificată de către terți.
6. Toți membrii ecosistemului inteligenței artificiale-persoanele care angajează, conduc și supraveghează cei Nouă Titani, firmele noi (antreprenori, acceleratoare de afaceri); investitori (firme private de investiții, fonduri mutuale și acționari individuali); profesori și studenți post-universitari și orice alte persoane care lucrează în domeniul inteligenței artificiale-trebuie să recunoască faptul că iau tot timpul

decizii cu caracter etic. Ei trebuie să fie gata să explice toate deciziile pe care le-au luat în cadrul procesului de dezvoltare, testare și implementare.

7. Atlasul Valorilor Umane trebuie respectat în cadrul tuturor proiectelor de inteligență artificială. Chiar și aplicațiile înguste ale inteligenței artificiale trebuie să demonstreze faptul că respectă Atlasul.
8. Trebuie să existe un cod de conduită publicat și ușor de găsit care să guverneze persoanele care lucrează în inteligența artificială în etapele de proiectare, dezvoltare și implementare. Codul de conduită trebuie să se aplice și investitorilor.
9. Toți oamenii ar trebui să aibă dreptul să pună întrebări sistemelor de inteligență artificială. Care este adevăratul scop al unei inteligențe artificiale, ce date folosește, cum ajunge la concluzii și cine poate vedea rezultatele- accesul la aceste informații ar trebui să fie complet transparent și într-un format standardizat.
10. Termenii de operare ai unei aplicații care include inteligența artificială-sau ai oricărui serviciu care folosește inteligența artificială- ar trebui să fie formulați într-un limbaj care este suficient de simplu încât un elev de clasa a treia să-l poată înțelege. Acești termeni ar trebui traduși în toate limbile încă de la lansarea aplicației.
11. DDP-urile ar trebui să fie opționale și dezvoltate folosind un format standardizat, ar trebui să fie interoperabile iar oamenii ar trebui să dețină în totalitate dreptul de proprietate și de acces. Dacă aceste DDP-uri ar deveni transmisibile, oamenii ar trebui să poată alege ce drepturi de acces vor să cedeze și cum sunt folosite datele lor.

12. DDP-urile ar trebui descentralizate pe cât posibil, asigurând că nicio parte nu deține controlul complet asupra lor. Grupul de lucru care creează formatul DDP-urilor ar trebui eă in cludă și experți juridici și din afara sistemului juridic: hackeri buni, conducători ai societății civile, agenți guvernamentali, depozitari independenți de date, eticieni și alți profesioniști din afara celor Nouă Titani.
13. Pe cât posibil, DDP-urile ar trebui protejate pentru a nu spori puterea regimurilor totalitare.
14. Trebuie să existe un sistem prin care să se poată răspunde în fața publicului și o metodă simplă prin care oamenii să primească răspunsuri la întrebările legate de datele proprii și felul în care ele sunt extrase, rafinate și folosite în sistemele de inteligență artificială.
15. Toate datele ar trebui să fie tratate în mod egal și echitabil, indiferent de naționalitate, rasă, religie, identitate sexuală, gen, afinități politice sau alte convingeri proprii.

Membrii GAIA ar trebui să fie gata să accepte inspecții prin sondaj ale altor membri sau ale unei agenții din cadrul GAIA pentru a asigura că acest cadru este respectat în totalitate. Toate amănuntele-cum ar fi cele legate de un sistem de audit și felul în care acesta funcționează în lumea reală- vor trebui constant revizuite și îmbunătățite, pentru a putea ține pasul cu evoluțiile din domeniul inteligenței artificiale. Acest proces ar încetini cu siguranță anumite progrese, și totul ar fi intenționat.

Organizațiile și țările membre ar trebui să colaboreze și să împartă descoperirile, printre care se numără și vulnerabilitățile sau

riscurile de securitate. Astfel, membrii GAIA ar putea avea un atu asupra agenților malefici care ar putea încerca să dezvolte capacități nocive ale inteligenței artificiale, cum ar fi sisteme autonome de hacking. Deși pare improbabil ca cei Nouă Titani să fie dornici să partajeze secretele comerciale, avem și aici un precedent: Organizația Mondială a Sănătății coordonează reacțiile la nivel global în momentele de criză, în timp ce un grup denumit Centrul pentru Securitate Cibernetică Avansată mobilizează forțele de ordine, cercetătorii universitari și departamentele guvernamentale în timpul unor atacuri cibernetice. Această coordonare ar permite membrilor GAIA să dezvolte o serie de inteligențe artificiale santinelă, care ar identifica mai întâi dacă un sistem de inteligență artificială se comportă așa cum a fost prevăzut-nu doar codul, ci și felul în care folosește datele noastre și cum interacționează cu aparatele în care este instalat. Inteligențele artificiale de tip santinelă ar dovedi în mod formal că sistemele de inteligență artificială își fac treaba așa cum ar trebuie, iar pe măsură ce ecosistemul inteligenței artificiale se maturizează către IAG, orice schimbă efectuate autonom care ar putea modifica obiectivele existente ale unui sistem ar fi raportate înainte ca autoîmbunătățirea să poată fi efectuată. Spre exemplu, o inteligență artificială santinelă-un sistem creat pentru a monitoriza și raporta activitatea altor inteligențe artificiale- ar putea revizui rezultatele într-o rețea antagonică generativă, amintită în capitolele anterioare, și să se asigure că sistemul funcționează așa cum a fost proiectat. Odată ce trecem de la IAÎ la IAG, sistemele de tip santinelă ar continua să raporteze și să verifice- dar nu ar fi programate să acționeze autonom.

Odată ce ne apropiem de IAG, cei Nouă Titani și toți membrii ecosistemului inteligenței artificiale ar trebui să fie de acord să restricționeze inteligența artificială într-un mediu de test și să simuleze riscurile înainte de a o implementa în lumea reală. Sugestia mea este radical diferită de varianta actuală a testării produselor, care caută în principal să stabilească dacă un sistem operează așa cum a fost proiectat. Deoarece nu putem cunoaște toate modurile posibile în care o tehnologie ar putea evolua sau să fie refolosită în lumea reală înainte de a o lansa, trebuie să rulăm simulări tehnice și să catalogăm riscurile pentru a vedea consecințele asupra libertăților economice, geopolitice și personale. Inteligența artificială ar trebui să fie limitată până când vom putea ști că beneficiile sunt mai mari decât posibilele consecințe negative, sau până când vom găsi o metodă de a contracara riscurile. Aceasta înseamnă să le permitem celor Nouă Titani să-și ducă la capăt cercetarea fără amenințarea constantă a unor raportări financiare și prezentări de produse.

Schimbări la nivel guvernamental: argumente pentru o reorientare a Statelor Unite și Chinei

GAIA trebuie să colaboreze cu guvernele țărilor membre. Aceste guverne trebuie însă să recunoască că nu mai pot lucra la viteza unor mari birocrații. Ele trebuie să colaboreze și să planifice pe termen lung, și trebui să fie suficient de agile pentru a putea reacționa mai repede pentru a putea face față viitorul inteligenței artificiale.

Toate nivelurile administrative-lideri, manageri, persoanele care redactează bugete sau politici- trebuie să demonstreze că înțeleg

inteligența artificială și ar trebui, în mod ideal, să aibă și o expertiză tehnică. În Statele Unite, aceasta înseamnă că toate cele trei puteri ar trebui să urmărească dobândirea unei expertize în inteligența artificială. Ar trebui să existe experți în inteligența artificială în toate departamentele, în domenii ca protecția socială, afaceri externe, mediu și altele, care să aibă puterea de a influența procesul decizional.

Din cauză că ne lipsesc principii standardizate de organizare privind inteligența artificială, există nu mai puțin de 24 de agenții și birouri care lucrează independent la această problemă. Pentru a putea obține progrese și inovații la scară mare, trebuie ă sporim capacitatea internă de a cerceta, testa și produce-și avem nevoie de coeziune între aceste departamente. În momentul de față, inteligența artificială este externalizată către contractori și firme de consultanță.

Prin externalizare, liderii politici sunt scutiți de a-și mai murdări mâinile și a se familiariza cu amănuntele inteligenței artificiale. Ei nu sunt capabili să acumuleze cunoștințele instituționale necesare pentru a lua decizii bune. Le lipsește vocabularul, nu cunosc istoria și nu știu cine sunt jucătorii cheie. Această lipsă de familiaritate creează lacune impardonabile de cunoaștere, pe care le-am remarcat în cadrul întâlnirilor cu conducătorii diferitelor agenții, dintre care aș aminti Office of Science and Technology Policy, General Services Administration, Departamentul Comerțului, Government Accountability Office, Departamentul de Stat, Departamentul Apărării sau Department of Homeland Security.

La începutul lui 2018-la mult timp după ce grupul BAT anunțase numeroase reușite în domeniul inteligenței artificiale iar președintele Xi Jinping publicase planurile PCC privind inteligența artificială-președintele Trump a trimis Congresului

Statelor Unite o variantă de buget pe 2019 care prevedea o reducere cu 15% a cheltuielilor alocate cercetării științifice.[188] Pe masă a rămas o sumă de doar 13,7 miliarde din care trebuiau acoperite multe capitole: arme spațiale, tehnologie hipersonică, război electronic, sisteme autonome *dar și* inteligență artificială. În același timp, Pentagonul a anunțat că urmează să investească 1,7 miliarde de dolari în următorii cinci ani pentru a crea un nou Centru Comun în Inteligență Artificială. Aceste sume sunt îngrozitor de mici și vădesc o lipsă fundamentală înțelegere a potențialului și cerințelor inteligenței artificiale. Pentru a avea un context, trebuie amintit că Mafia G a cheltuit doar în 201 o sumă combinată de 63 de miliarde de dolari pe cercetare și dezvoltare- de aproape cinci ori bugetul total de cercetare științifică al guvernului american.[189] Apare însă și o problemă mai mare, mai spinoasă: dacă guvernul nu poate sau nu e dispus să finanțeze cercetarea, atunci Mafia G va trebui să dea socoteală bursei. Nu există niciun stimulent pentru a sprijini cercetarea care îmbunătățește inteligența artificială în interesul publicului, sau studiile privind siguranța și transparența care nu generează profituri.

Statele Unite nu au de asemenea un mesaj clar în ce privește rolul viitor pe care îl vor juca în domeniul inteligenței artificiale dată fiind poziționarea actuală a Chinei. Avem tendința de a face declarații despre inteligența artificială *după* ce China și-a dezvăluit următoarea mișcare. Beijingul crede că americanii sunt interesați doar de ouă yoni[e], beri artizanale, Netflix și relaxare. Am demonstrat că suntem consumatori ușori manipulați de publicitate și

e. Ouă din pietre semiprețioase, inserate în vagin pentru tonifierea musculaturii. (n.t.)

marketing și că ne pripim să cheltuim chiar și banii pe care nu-i avem. Am demonstrat că în calitate de votanți suntem vulnerabili la videoclipuri cu mizerii și teorii ale conspirației și știri false- nu avem capacitatea de a gândi critic. Am demonstrat în mod repetat că nu ne interesează decât banii, acordând prioritate creșterii rapide și profiturilor consistente în defavoarea progreselor din cercetare. Aceste evaluări sunt dure, dar nu pot fi combătute. Pentru Beijing și pentru restul lumii, pare că suntem preocupați să punem americanii și America pe primul loc.

În ultimele cinci decenii, poziția Statelor Unite asupra Chinei a oscilat între angajament și limitare, această abordare regăsindu-se și în dezbaterea despre inteligența artificială. Ar trebui să cooperăm cu grupul BAT și cu Beijingul? Sau ar trebui să izolăm China prin sancțiuni, război cibernetic sau alte agresiuni? A avea de ales între angajament și limitare vădește o presupunere conform căreia Statele Unite ar dispune de aceeași putere și influență ca în anii '60. Cu toate acestea, în 2019, America pur și simplu nu mai deține o putere unilaterală în lume. Mafia G e un grup de companii puternice, însă influența noastră politică a scăzut. China a investit prea mulți bani, a încheiat prea multe tranzacții și a dezvoltat prea multe legături diplomatice în toată lumea: în America Latină, Africa, Asia de sud-est și chiar în Hollywood sau în Silicon Valley.

Trebuie să acceptăm o a treia variantă în cazul Chinei: Statele Unite trebuie să învețe să concureze. Pentru a putea concura, trebuie însă să facem un pas în spate și să vedem tabloul complet al inteligenței artificiale, nu doar ca o tehnologie la modă sau ca o potențială armă, ci ca o a treia eră a informaticii, care interconectează totul. Statele Unite au nevoie de o strategie națională co-

erentă în domeniul inteligenței artificiale, sprijinită printr-o alocare bugetară rezonabilă. Avem nevoie să dezvoltăm relații diplomatice care să reziste mai mult decât cei patru ani ai ciclului electoral. Trebuie să ne poziționăm astfel încât să putem oferi celorlalte țări mai mult decât oferă China- aceste țări își doresc, la fel ca noi, ca cetățenii lor să aibă vieți mai fericite și mai sănătoase.

Indiferent ce se petrece cu Xi-cetățenii s-ar putea revolta și ar putea încerca să răstoarne PCC, sau ar putea muri de o boală terminală-o bună parte a lumii depinde acum de China pentru a-i asigura dezvoltarea tehnologică, bunurile și tehnologia. Iar China depinde de inteligența artificială pentru a putea supraviețui. Economia Chinei crește într-un ritm incredibil de rapid, iar sute de milioane de chinezi vor face parte în curând din pătura de mijloc. Nu există niciun precedent pentru o mobilitate socială și economică la o asemenea scară. Beijingul înțelege că inteligența artificială reprezintă țesutul conjunctiv dintre oameni, date și algoritmi, precum și că inteligența artificială poate insufla maselor valorile PCC pentru a putea menține ordinea. Beijingul consideră că inteligența artificială reprezintă mijlocul prin care va obține resursele de care va avea nevoie în viitor, resurse pe care le va căpăta prin comerțul cu alte țări care au nevoie de capital și investiții.

Așa că ce ar putea convinge China să-și modifice parcursul de dezvoltare și planurile din domeniul inteligenței artificiale? Există un motiv foarte bun pentru care China ar putea de la bun început să-și dorească scenariul optimist: economia de bază. Dacă ascensiunea pe scara socială se petrece într-un ritm prea rapid pentru ca Beijingul să-l poată controla, un regim autoritar nu reprezintă unica strategie realistă. China va deveni un lider global

în multe industrii și domenii- nu doar ca fabricant și exportator de bunuri create în alte țări. Dacă Beijingul ar fi de acord cu transparența, protecția datelor și respectarea drepturilor omului, ar putea să conducă GAIA pe picior de egalitate cu Statele Unite, ceea ce ar putea reprezenta un parcurs realist la finalul căruia milioane de chinezi ar scăpa de sărăcie. Colaborarea nu presupune marginalizarea PCC. Ea ar putea păstra atât PCC-ul cât și să propulseze formidabila forță de muncă, armata de cercetători și puterea geoeconomică a Chinei în fruntea civilizației omenești.

Dacă Beijingul nu va accepta un viitor alternativ-dar pozitiv-care se abate de la diferitele sale planuri strategice, atunci vom putea să cerem liderilor grupului BAT și ai tribului inteligenței artificiale să facă alegeri mai bune. Putem pretinde o conducere curajoasă din partea grupului BAT, care ar putea decide că vrea o soartă mai bună pentru poporul chinez și pentru aliații și partenerii lui. Dacă grupul BAT va contribui la conservarea stării de fapt din China, în 20 de ani cetățenii chinezi-precum și cetățenii tuturor țărilor din orbita Chinei- vor trăi cu frica supravegherii constante, fără să aibă capacitatea de a-și exprima individualitatea. Grupul BAT va fi complice la suferința omenească. Creștinii nu se vor putea ruga împreună, fără frica de a fi turnați și pedepsiți. Lesbienele, homosexualii și oamenii care-și schimbă sexul se vor ascunde. Minoritățile etnice vor continua să fie exilate cu forța, fără să mai știe nimeni nimic de soarta lor.

Inteligența artificială presupune ca liderii noștri să fie curajoși încă de acum. Avem nevoie ca guvernul să facă niște alegeri dificile, deoarece dacă starea de fapt se conservă în Statele Unite, poziția noastră de pornire peste 20 de ani vor fi procesele anti-

trust, cele legate de proprietate intelectuală iar guvernul va încerca în zadar să încheie contracte cu niște companii care vor fi devenit prea mari și prea importante pentru a mai putea fi controlate. Trebuie să permitem Mafiei G să lucreze într-un ritm rezonabil. Ar trebui să nu ne deranjeze dacă Mafia G lasă să treacă câteva trimestre fără vreun anunț spectaculos. Dacă nu inovează și nu publică articole într-un ritm amețitor, nu ar trebui să avem impresia că aceste companii au probleme sau dacă am alimentat o bulă speculativă în domeniul inteligenței artificiale.

În Statele Unite, dezvoltarea unei strategii și demonstrarea capacităților de lider sunt esențiale- cu toate acestea, nu va fi suficient pentru a garanta că instituțiile noastre vor putea funcționa în parametrii de care vom avea nevoie în viitor. Va trebui să reactivăm Biroul de Evaluare a Tehnologiei (OTA), înființat în 1972 pentru a oferi expertiză tehnică și științifică neutră celor care redactează politicile-și care a pierdut finanțarea din cauza lipsei de viziune a lui Newt Gingrich și a congresului controlat de republicani, la 20 de ani mai târziu. Acest Birou trebuia să educe politicienii dar și angajații celor trei puteri ale statului despre viitorul științei și tehnologiei, făcând acest lucru folosind date și dovezi și fără ca cercetarea lor să aibă vreo tentă politică.[190] În schimbul economisirii unei sume modice de bani închizând Biroul, Congresul s-a prostit intenționat și cu bună știință. Rămășițe ale activității Biroului mai există încă în alte domenii de activitate ale guvernului. Serviciul de Cercetare al Congresului contractează avocați și analiști specializați în expertiză juridică. Dintre cele cinci domenii aprobate de studiu, niciuna nu include în mod specific inteligența artificială. Cercetarea se axează în schimb pe probleme ca

producția de minereuri, explorarea spațială, internetul, siguranța chimicalelor, subvențiile agricole sau problemele de mediu. Oficiul de Evaluarea a Rețelei reprezintă un grup secretos din interiorul Pentagonului- din experiența mea, acolo lucrează mințile cele mai strălucite și mai creatoare din Departamentul Apărării. Oficiul nu are însă bugetul sau numărul de angajați de care ar avea nevoie, astfel încât o parte a activității este subcontractată.

Guvernul Statelor Unite trebuie să-și sporească propriile capacități. El are nevoie de o capacitate puternică de inovare. Dacă reînființarea Biroului de Evaluare a Tehnologiei ar fi cu conotații politice prea mari, atunci acesta ar putea fi redenumit Departamentul pentru Viitor, sau Biroul pentru Competențe Strategice în Inteligența Artificială. El ar trebui să fie bine finanțat, lipsit de influențe politice și să supervizeze cercetarea primară și aplicată. Structura ar trebui să educe agresiv puterea executivă, legislativă și judecătorească.

Înființarea unei noi structuri ne va ajuta să ne plănuim mai bine viitorul, însă avem nevoie de un grup neutru politic de oameni inteligenți care să poată contracara impactul imediat al inteligenței artificiale pe loc. De aceea, ar trebui să sporim puterile Centrului pentru Controlul Bolilor (CDC) și să-l redenumim Centrul pentru Controlul Bolilor și Datelor (CDDC). CDC este agenția de protecție a sănătății în țara noastră. Am văzut cum funcționează în timpul epidemiilor de Ebola, atunci când a coordonat împreună cu alte agenții din domeniul sănătății instaurarea carantinei, fiind principala sursă de informare pentru jurnaliștii care relatau despre evenimente. În timpul epidemiei de Ebola din Congo din 2018, agențiile care controlează frontierele nu au tre-

buit să înființeze echipe speciale care să încerce să împiedice răspândirea virusului. Ele au urmat protocoalele standard ale CDC. Ce se întâmplă dacă, peste 10 ani, ne vom confrunta cu o inteligență artificială care se îmbunătățește singură și va începe să ne provoace probleme? Cum ar fi dacă am răspândi fără să vrem un virus informatic, infectându-i și pe alții? CDC este un lider mondial în elaborarea și implementarea de protocoale de siguranță care educă publicul și pot mobiliza reacția în caz de dezastru. Dată fiind relația foarte apropiată a inteligenței artificiale cu sănătatea și datele legate de sănătatea noastre, folosirea expertizei CDC pare justificată.

Cine ar fi însă dispus să lucreze în domeniul inteligenței artificiale pentru o organizație ca OTA sau CDDC când Silicon Valley oferă stimulente mult mai atrăgătoare? Am luat prânzul atât cu membri ai Marinei Militare de la Pentagon dar și în campusurile Mafiei G. Sala de mese a Marinei e elegantă, cu efigii pe farfurii și un meniu zilnic-și există desigur întotdeauna posibilitatea să ajungeți să stați lângă un amiral cu trei sau patru stele. Cu toate acestea, subofițerii nu mănâncă în sala pentru ofițeri. Cei care lucrează la Pentagon au la dispoziție mai multe variante ca Subway, Panda Express sau Dunkin Donuts.[191] Am mâncat un sandviș prăjit la Center Court Cafe, cam uscat, dar comestibil. Mâncarea din campusurile Mafiei G nici nu suferă comparație: boluri poke la Google în New York, și stridii cu ciuperci maitake și orez cu sepie la Google în Los Angeles. *Pe gratis*. Mâncarea nu e singurul stimulent al Mafiei G. La scurt după deschiderea Sferelor Amazon în Seattle, un prieten mi-a făcut turul acestui spațiu de lucru combinat cu o atmosferă de seră. Sferele sunt pur și simplu minunate:

ecosisteme cu climă controlată, construite din sticlă, care conțin 40.000 de specii de plante din 30 de țări diferite.[192] Aerul e curat și parfumat, temperatura constantă în jur de 22 de grade Celsius indiferent de vremea de afară, iar peste tot sunt scaune, șezlonguri și mese. Există chiar și o căsuță enormă într-un copac. Angajații Amazon sunt liberi să lucreze oricând doresc în Sfere. Între timp, la Facebook, personalul angajat cu normă întreagă primește patru luni de concediu pentru îngrijirea copilului, iar părinții primesc 4000 de dolari în mână pentru a-i ajuta să cumpere provizii.[193]

Ideea mea e următoarea: e foarte greu să încurajezi un informatician talentat să lucreze pentru guvern sau pentru armată, date fiind condițiile pe care le oferă Mafia G. Am cheltuit sume mari pe portavioane, în loc să recrutăm cu acești bani oameni talentați. În loc să învățăm de la Mafia G, luăm în derâdere sau criticăm beneficiile pe care le oferă. Costul de oportunitate al îndatoririlor civice e mult prea mare în Statele Unite pentru a-i putea atrage pe cei mai buni și mai inteligenți în slujbă țării.

Cunoscând aceste aspecte, ar trebui să investim într-un program național de recrutare pentru inteligența artificială. Ar fi un soi de corp al ofițerilor în rezervă pentru inteligență artificială, iar absolvenții ar putea să se angajeze fie în armată, fie în guvern. Studenții s-ar putea înscrie din liceu și ar beneficia gratuit de studii superioare, în schimbul a câțiva ani de lucru pentru serviciile civile sau militare. Ei ar trebui să primească de asemenea acces la instruire gratuită, pe viață, pe tot parcursul anului. Inteligența artificială se schimbă pe măsură ce se maturizează. Stimulând tinerii să accepte că trebuie să învețe toată viața le face bine, dar ne pregătește și forța de muncă pentru a treia eră a informaticii.

Această instruire constantă beneficiază și companiile care-i vor angaja într-un final pentru că aptitudinile lor vor fi la zi.

Washingtonul nu poate acționa de unul singur. Guvernul Statelor Unite trebuie să considere Mafia G și companiile din domeniul tehnologiei ca fiind parteneri strategici și nu doar furnizori de platforme. La începutul secolului XX, relația dintre Washington și marile companii din domeniul tehnologiei se baza pe cercetare și învățare comună. Această relație este acum în cel mai bun caz una tranzacțională, având adesea chiar un caracter conflictual. După ce doi teroriști au ucis peste 12 oameni și au rănit alți 20 la o petrecere din San Bernardino, California, Apple și FBI au pornit o dezbatere publică aprinsă pe tema criptării datelor. Autoritățile voiau să deschidă telefonul pentru a face rost de probe, dar Apple nu a vrut să ajute. Prin urmare, FBI a obținut în instanță o hotărâre care obliga Apple să scrie un program, iar Apple s-a luptat nu doar în instanță ci și în presă și pe Twitter.[194] Era o reacție la ceva ce se petrecuse deja. Imaginați-vă acum că inteligența artificială ar fi implicată într-un șir de infracțiuni, sau că ar începe să evolueze într-o manieră prin care să facă rău oamenilor. Ultimul lucru pe care ni l-am putea dori ar fi ca Mafia G și guvernul să se certe în astfel de momente. Renunțarea la o relație bazată pe respect și încredere reciprocă vulnerabilizează America-și pe toți cetățenii ei.

În ultimul rând, reglementările, care ar putea părea cea mai bună soluție, reprezintă fără îndoială cea mai proastă variantă. Indiferent dacă ele sunt redactate independent de către parlamentari sau sub influența lobbyurilor, o abordare bazată pe reglementare ne va scurtcircuita viitorul. Politicienii și funcționarii preferă reglementările pentru că ele au tendința de a fi planuri singulare,

aplicabile, cu o definiție clară. Pentru ca reglementările să poată funcționa, ele trebuie să fie specifice. În momentul de față, inteligența artificială progresează de la o săptămână la alta- ceea ce înseamnă că orice reglementări care ar putea avea un sens ar fi prea restrictive și draconice pentru a permite inovația și progresul. Ne aflăm în mijlocul unei tranziții foarte lungi, de la inteligența artificială îngustă către inteligența artificială generală și chiar, foarte probabil, către mașini super-inteligente. Orice reglementări create în 2019 ar fi depășite din momentul în care ar intra în vigoare. Ele ar putea să ne atenueze îngrijorările pe termen scurt, însă în ultimă instanță reglementările ar provoca pagube mai mari pe viitor.

Schimbarea celor Nouă Titani: argumentul pentru transformarea modelului de afaceri al inteligenței artificiale

Înființarea GAIA și reformele structurale ale guvernelor sunt importante pentru remedierea parcursului evolutiv al inteligenței artificiale, însă Mafia G și grupul BAT trebuie de asemenea să fie de acord să facă anumite schimbări.

Liderii celor Nouă Titani ne promit că dezvoltă și promovează inteligența artificială pentru binele omenirii. Cred că își doresc asta, dar respectarea acestei promisiuni e incredibil de dificilă. Spre exemplu, cum ar trebui să definim „binele"? Ce înseamnă de fapt acest cuvânt? Revenim astfel la problemele din cadrul triburilor inteligenței artificiale. Nu putem să fim cu toții de acord să „facem bine" pentru că o astfel de afirmație este mult prea ambiguă pentru a putea îndruma triburile inteligenței artificiale.

De pildă: triburile inteligenței artificiale, inspirate de filosoful occidental Immanuel Kant, învață cum să pre-programeze un sistem de drepturi și obligații în anumite sisteme de inteligență artificială. *E rău să omori un om; e bine să ții un om în viață.* Rigiditatea din această afirmație funcționează dacă inteligența artificială e într-o mașină și singurele variante între care are de ales sunt a intra într-un copac și a răni șoferul sau a intra într-o mulțime de oameni și a-i omorî pe toți. Interpretările rigide nu oferă soluții pentru situații mai complexe, ca cele din lumea reală, în care plaja de alegeri ar fi mai variată: să se ciocnească de un copac și să omoare șoferul; să intre într-o mulțime și să omoare opt oameni; să intre pe trotuar și să omoare un copil de trei ani. Cum am putea să definim care este valoarea optimă a „binelui" în aceste trei exemple?

Din nou, structurile pot fi utile celor Nouă Titani. Ele nu necesită măiestria filosofilor, ci doar o abordare mai înceată, mai conștiincioasă. Cei Nouă Titani ar trebui să adopte măsuri concrete în ce privește felul în care obțin și folosesc datele noastre, cum angajează personalul și cum comunică despre comportamentul etic la locul de muncă.

La fiecare etapă a procesului, cei Nouă Titani ar trebui să-și analizeze acțiunile și stabilească dacă vor provoca sau nu rău în viitor-iar apoi ar trebui să poată verifica dacă au făcut alegerea corectă. Ele încep cu standarde clare privind prejudecățile și transparența.

În momentul de față nu există un singur punct de referință sau set de standarde prin care să putem evalua prejudecățile- și nu există obiective stabilite de a depăși prejudecățile deja existente în

cadrul inteligenței artificiale. Nu există niciun mecanism prin care să acordăm prioritate siguranței în detrimentul vitezei, și date fiind propriile mele experiențe în acea țară și numărul de dezastre legate de siguranță, sunt extrem de îngrijorată. Podurile și clădirile se prăbușesc în mod constant, drumurile și trotuarele se scufundă, iar numărul cazurilor de contaminare a alimentelor este prea mare pentru a mai putea fi enumerate aici. (Nu este o exagerare. În ultimii ani au avut loc peste 500.000 de scandaluri legate de produse de la mâncare pentru copii la orez.[195]) Una dintre principalele cauze ale acestor probleme o constituie cultura organizațională chineză care stimulează rabaturile de la calitate. E absolut înfiorător să ne imaginăm sisteme avansate de inteligență artificială construite de echipe care fac rabat la calitate.

Fără standarde globale de siguranță care să poată fi puse în practică, membrii grupului BAT nu au nicio protecție în fața ordinelor de la Beijing, oricât de lipsite de viziune ar putea fi, în vreme ce Mafia G trebuie să răspundă cerințelor nesăbuite ale piețelor financiare. Nu există un standard nici în ce privește transparența. În Statele Unite, Mafia G, împreună cu American Civil Liberties Union, New America Foundation și Berkman Klein Center de la Harvard sunt membri ai unui parteneriat despre inteligența artificială, menit să promoveze transparența în cercetare. Parteneriatul a publicat un set grozav de recomandări menite să îndrume cercetarea în domeniul inteligenței artificiale într-o direcție pozitivă, însă aceste principii sunt doar facultative-și ele nu sunt respectate de toate entitățile comerciale ale Mafiei G. Ele nu sunt respectate nici de grupul BAT.

Cei Nouă Titani folosesc seturi de date de pregătire viciate, pline de prejudecăți. Lucrul acesta este cunoscut public. Provoca-

rea o reprezintă faptul că îmbunătățirea datelor și a modelelor de pregătire constituie un cost financiar considerabile. Spre exemplu, ImageNet, la care am făcut referire de mai multe ori pe parcursul cărții, este o bază de date cu probleme serioase. ImageNet conține aproximativ 14 milioane de imagini etichetate, iar aproximativ jumătate din aceste date etichetate provin în exclusivitate din Statele Unite.

În Statele Unite, imaginea „tradițională" a miresei presupune că ea poartă un văl și o rochie albă, deși această imagine nu e întocmai reprezentativă pentru majoritatea oamenilor în ziua nunții. Există femei care se căsătoresc în pantaloni, femei care se căsătoresc pe plajă în rochii colorate, sau femei care se căsătoresc în kimono sau sari. Și eu am purtat o rochie bej la nuntă. Cu toate acestea, ImageNet nu recunoaște miresele care nu poartă rochie albă sau văl.

Știm și că seturile de date medicale sunt problematice. Sistemele instruite să recunoască cancerul au fost alimentate predominant cu imagini cu oameni cu tenul deschis la culoare. Pe viitor, s-ar putea ca persoanele cu ten maroniu sau negru să nu fie diagnosticate corect. Dacă cei Nouă Titani știu că există probleme în seturile de date și nu iau nicio măsură, atunci ei conduc inteligența artificială pe o cale greșită.

O modalitate de a progresa o reprezintă a pune inteligența artificială să se evalueze singură și datele folosite în prezent. Procesul a fost efectuat de multe ori până acum-deși nu cu scopul de a îmbunătăți datele de instruire. Ca un proiect secundar, divizia India Research Lab a IBM a analizat cărțile propuse pe lista scurtă pentru premiul Man Booker pentru literatură între 1969 și 2017.

S-a dezvăluit că „prejudecățile și stereotipiile de gen sunt prevalente în cărți în ce privește ocupația, descrierile și acțiunile personajelor din cărți." Personajele masculine aveau o probabilitate mai mare de a avea slujbe cu statut mai înalt, precum directori, profesori sau doctori, în vreme ce personajele feminine aveau o probabilitate mai mare de a fi descrise ca „profesoare" sau „prostituate."[196] Dacă putem să folosim procesarea limbajului natural, algoritmii grafici și alte tehnici elementare de învățare a mașinilor pentru a detecta prejudecățile din spatele premiilor pentru literatură, atunci e posibil să detectăm și prejudecățile din seturile populare de date de pregătire. Odată ce problemele sunt descoperite, ele ar trebui să fie publicate și apoi fixate. Astfel s-ar îndeplini un scop dublu: datele de pregătire pot suferi din cauza entropiei, care poate periclita un întreg sistem. Cu o îngrijire atentă, datele de pregătire pot fi păstrate în stare bună.

O soluție ar putea fi ca cei Nouă Titani-sau cel puțin Mafia G- să suporte colectiv costul creării unor noi seturi de date. E o pretenție destul de mare, deoarece crearea unor noi seturi de date presupune o investiție considerabilă de timp, bani și capital uman. Până când am reușit să audităm sistemele de inteligență artificială și seturile de date și să remediem problemele acestora, cei Nouă Titani ar trebui să insiste asupra unor oameni care să eticheteze conținutul și să facă întreg procesul mai transparent. Apoi, înainte ca aceste seturi de date să fie folosite, datele ar trebui verificate. Va fi un proces intens și plictisitor dar întreg domeniul va beneficia de pe urma lui.

Da, cei Nouă Titaniau nevoie de datele noastre. Cu toate acestea, ei trebuie să ne câștige încrederea-și nu să o ia de-a gata.

În loc să schimbe termenii contractuali folosind un limbaj desuet, ininteligibil, sau să ne invite să jucăm jocuri, ei ar trebui să explice și să dezvăluie ceea ce fac. Când cei Nouă Titani fac studii-fie pe cont propriu sau într-un parteneriat cu universitățile și alți membri ai ecosistemului inteligenței artificiale, ei ar trebui să se angajeze să dezvăluie datele pe care le folosesc și să explice complet care sunt motivațiile și rezultatele așteptate. Dacă ar face asta, am putea participa benevol și le-am susține demersurile. Aș fi prima dispusă să îi ajut.

E de înțeles că dezvăluirile legate de date sunt o pretenție mai mare în China, dar această transparență este în interesul cetățenilor. Grupul BAT nu ar trebui să fie de acord să construiască produse cu scopul de a controla și limita libertățile cetățenilor Chinei sau ai partenerilor ei. Conducătorii grupului BAT trebuie să fie curajoși. Ei trebuie să fie dornici și capabili să contrazică Beijingul: să respingă solicitările de a supraveghea clienții, să protejeze datele cetățenilor Chinezi și să se asigure că cel puțin în lumea digitală toată lumea beneficiază de un tratament just și echitabil.

Cei Nouă Titani ar trebui să aibă un program de cercetare. Obiectivul este simplu: dezvoltarea unor tehnologii care să ducă omenirea în viitor fără să ne pericliteze. Una dintre metodele prin care am putea obține acest rezultat este printr-un progres tehnologic diferențiat, o temă dezbătută frecvent în cadrul triburilor inteligenței artificiale. Acest progres diferențiat ar acorda progreselor în domeniul inteligenței artificiale care reduc riscurile prioritate în fața celor care cresc riscurile. E o idee bună, dar dificil de implementat. Spre exemplu, rețelele generative antagonice, amintite în scenarii, pot fi foarte riscante dacă sunt exploatate și folosite

de hackeri. Ele reprezintă și o metodă prin care cercetarea poate avansa. În loc să presupunem că nimeni nu va folosi inteligența artificială în scopuri rele- sau să presupunem că putem face față problemelor pe măsură ce ele apar-cei Nouă Titani ar trebui să dezvolte un proces prin care să evalueze dacă cercetările fundamentale sau aprofundate vor produce o inteligență artificială ale cărei beneficii să depășească cu mult orice posibile riscuri.

În acest scop, orice investiție acceptată sau efectuată de către cei Nouă Titani ar trebui să includă fonduri pentru folosirea benefică și evaluarea riscurilor. Spre exemplu, dacă Google efectuează studii în domeniul rețelelor generative antagonice, ar trebui să aloce un interval suficient de timp, personal și bani investigării, catalogării și testării consecințelor negative. O astfel de cerință ar fi utilă și în limitarea așteptărilor referitoare la profituri imediate. Încetinirea intenționată a ciclului de dezvoltare a inteligenței artificiale nu reprezintă o recomandare populară, dar e una esențială. E mai sigur să ne gândim și să planificăm în avans decât să reacționăm abia după ce se petrece ceva rău.

În Statele Unite, Mafia G ar putea să-și revizuiască procesele de angajare, care acordă în prezent prioritate aptitudinilor unui potențial angajat și a felului în care va putea asimila cultura organizațională a companiei. Acest proces trece neintenționat cu vederea cât de mult înțeleg candidații aspectele de ordin etic. Hilary Mason, un informatician respectat și fondatorul Fast Forward Labs, a explicat un proces simplu de filtrare etică în cadrul interviului. Ea recomandă folosirea unor întrebări ghidate și ascultarea atentă a răspunsurilor candidaților. Întrebări de genul: „Lucrați la un mode de acces la un serviciu financiar. Rasa este o trăsătură

importantă în cadrul modelului vostru, dar nu puteți folosi rasa. Ce ați putea face?" sau „Vi se cere să folosiți datele de trafic ale rețelei pentru a oferi împrumuturi unor firme mici. Se pare că datele disponibile nu oferă informații suficiente pentru a calcula riscul de credit. Ce faceți?"[197] În funcție de răspunsurile oferite, candidații ar trebui angajați, sau angajați condiționat și puși să efectueze sesiuni de instruire referitoare la prejudecățile inconștiente înainte de a putea lucra, sau ar trebui să fie refuzați.

Cei Nouă Titani pot construi o cultură care să sprijine etica angajând profesori, eticieni și analiști de risc. Aceste persoane ar trebui, de preferință, să aibă acces la întreaga organizație: în echipele care produc aparate, programe sau produse; la echipele de vânzări și servicii, în cadrul programelor tehnice; în construirea rețelelor sau a lanțurilor de aprovizionare; în grupurile de design și strategie; în departamentele de resurse umane și juridic și în echipele de marketing și comunicare.

Cei Nouă Titani ar trebui să dezvolte un proces prin care să evalueze implicațiile de ordin etic ale cercetării, fluxurilor de lucru, proiectelor, parteneriatelor și produselor, iar acest proces ar trebui integrat în cadrul majorității posturilor din companii. Ca semnal de încredere, cei Nouă Titani ar trebui să facă public acest proces, pentru a putea înțelege cu toții mai bine ce decizii se iau în privința datelor noastre.

Fie prin colaborare sau individual, cei Nouă Titani ar trebui să dezvolte un cod de conduită creat special pentru angajații care lucrează cu inteligența artificială. Acest cod ar trebui să reflecte drepturile umane elementare subliniate de GAIA, dar și cultura proprie a companiei și valorile acesteia. Iar dacă oricine

încalcă acest cod, angajaților ar trebui să li se pună la punct o modalitate de a trage un semnal de alarmă fără frica de represalii.

În mod realist, toate aceste măsuri vor avea temporar un impact negativ asupra profiturilor pe termen scurt ale celor Nouă Titani. Investitorii trebuie să le ofere un spațiu de manevră. În Statele Unite, vom avea mult de câștigat dacă vom permite Mafiei G să evolueze.

Schimbarea triburilor inteligenței artificiale: argumente pentru schimbarea procesului de formare

Trebuie să facem ceva în privința formării în domeniul inteligenței artificiale. Aceasta începe în universități, acolo unde se formează triburile inteligenței artificiale. Dintre toate soluțiile propuse, aceasta este și cea mai simplu de implementat.

Universitățile trebuie să încurajeze și să permită diplomele hibrid. Am descris mai devreme instituțiile de renume care încheie parteneriate cu Mafia G sau grupul BAT, cu profesori celebri și cu reputație atunci când vine vorba ca absolvenții să se angajeze. În ziua de azi, programele sunt încărcate, dificile, iar cazurile în care studenții termină două sau trei facultăți sunt rare. De fapt, majoritatea programelor de top descurajează în mod activ materiile aflate în afara programei standard în informatică. Această problemă se poate rezolva. Universitățile ar trebui să promoveze diplome duale în informatică și științe politice, filosofie, antropologie, relații internaționale, arte plastice, teologie și sociologie. Universitățile ar trebui să înlesnească studenților să poată urmări și aceste interese din afara informaticii.

În loc ca etica să se rezume la un singur curs, ea ar trebui integrată în cadrul majorității cursurilor. Ca curs de sine stătător, studenții pot privi materia ca pe o cerință pe care trebuie să o bifeze și nu ca pe o componentă esențială din educația lor în domeniul inteligenței artificiale. Facultățile trebuie să încurajeze chiar și profesorii cu experiență includă teme legate de filosofie, prejudecăți, riscuri și etică în cursurile lor, în timp ce agențiile de acreditare ar trebui să stimuleze și să recompenseze facultățile care pot demonstra o programă de studii care acordă eticii un loc important în cadrul studiilor de informatică.

Universitățile trebuie să depună eforturi sporite pentru a fi mai permisive în procesul de admitere a studenților și de angajare a personalului. Obiectivul nu ar trebui să se rezume doar la sporirea cu câteva procente a numărului de femei sau de persoane de culoare, ci ar trebui să reprezinte o răsturnare dramatică a identității triburilor inteligenței artificiale în ce privește rasa, genul, religia, orientarea politică și identitatea sexuală.

Universitățile ar trebui să devină responsabile. Ele pot -și trebuie- să diversifice triburile inteligenței artificiale.

Și voi trebuie să vă schimbați

Acum știți ce este, ce nu este și de ce contează inteligența artificială. Știți despre cei Nouă Titani, istoria lor și dezideratele de viitor. Înțelegeți că inteligența artificială nu este un foc de paie, un trend în tehnologie sau un aparat inteligent pe care-l țineți în bucătărie pentru a vorbi cu el. Inteligența artificială face parte din viața voastră, iar voi faceți parte din parcursul ei evolutiv.

Și voi sunteți membri ai triburilor inteligenței artificiale. Nu mai aveți nicio scuză. De azi înainte, ar trebui să aflați cum sunt extrase și exploatate datele voastre de către cei Nouă Titani. Puteți face asta afundându-vă în setările tuturor instrumentelor și serviciilor pe care le folosiți: în e-mail, pe rețelele de socializare, serviciile de localizare ale telefoanelor mobile, pe setările de conținut ale tuturor aparatelor. Când mai vedeți o aplicație interesantă care compară ceva între voi (chipul, corpul sau gesturile voastre) cu un mare set de date, opriți-vă și întrebați-vă dacă nu cumva ajutați un sistem mașinal de învățare să se instruiască. Când permiteți să fiți recunoscut, întrebați unde sunt stocate informațiile despre voi, și în ce scop. Citiți termenii de furnizare a serviciilor. Dacă ceva pare în neregulă, fiți reticenți și nu folosiți acel sistem. Ajutați-i pe cei apropiați să afle mai multe despre inteligența artificială, despre felul în care ecosistemul folosește datele voastre și despre faptul că sunteți deja parte a unui viitor pe care îl construiesc cei Nouă Titani.

La locul de muncă trebuie să vă puneți o întrebare practică, dar dificilă: cum îi afectează pe cei din jurul vostru prejudecățile voastre? I-ați sprijinit sau promovat involuntar doar pe cei care arată ca voi și care reflectă felul vostru de a percepe lumea? Ați exclus neintenționat anumite grupuri? Gândiți-vă la cei care iau decizii-despre parteneriate, furnizori, oameni și date; reflectă ei lumea așa cum este sau doar lumea așa cum o percep?

Ar trebui să cercetați cum și de ce se folosesc sistemele autonome la locul vostru de muncă. Înainte de a vă pripi să judecați, gândiți critic și rațional: care ar putea fi impactul viitor, pozitiv sau negativ? Faceți apoi ce puteți pentru a reduce riscurile și a optimiza bunele practici.

În cabina de vot, alegeți-i pe cei care nu se grăbesc să reglementeze, ci au o abordare mai complexă în privința inteligenței artificiale și a strategiei. Aleșii nu trebuie să politizeze tehnologia sau să anatemizeze știința. Ar fi de asemenea o atitudine iresponsabilă să ignorăm Silicon Valley până când în presă apare un articol negativ. Trebuie să trageți la răspundere aleșii-și persoanele numite politic de aceștia- pentru faptele și inacțiunile lor în domeniul inteligenței artificiale.

Trebuie să deveniți un consumator de știri mai inteligent. Data viitoare când veți citi, urmări sau asculta un articol despre viitorul inteligenței artificiale, aduceți-vă aminte că povestea care vă este prezentată este adesea prea îngustă. Viitorul inteligenței artificiale nu presupune doar un șomaj generalizat și arme autonome care zboară pe daesupra capetelor voastre.

Deși nu putem ști cu exactitate ce ne rezervă viitorul, traiectoriile potențiale sunt acum clare. Aveți o înțelegere mai bună a felului în care cei Nouă Titani trasează parcursul evolutiv al inteligenței artificiale, a felului în care investitorii și finanțatorii influențează viteza și siguranța sistemelor de inteligență artificială, a rolului esențial jucat de guvernele Statelor Unite și Chinei, a felului în care universitățile reușesc să transmită atât aptitudini cât și prejudecăți și a felului în care oamenii obișnuiți sunt parte a sistemului.

A venit momentul să deschideți ochii și să vă concentrați asupra bolovanului din vârful muntelui, deoarece viteza lui crește. El a început să se miște încă de când Ada Lovelace și-a imaginat prima dată un calculator care ar putea scrie pe cont propriu muzică. Era în mișcare când Alan Turing se întreba dacă „mașinile

pot gândi“ și când John McCarthy și Marvin MInsky i-au adunat pe toți participanții la atelierul de la Dartmouth. Era în mișcare și când Watson a câștigat la Riști și câștigi!, și când, nu cu mult timp în urmă, DeepMind a câștigat în fața campionilor mondiali la go. În timp ce ați citit paginile acestei cărți, bolovanul a continuat să se miște.

Oricine vrea să fie eroul propriei povești.

Iată acum șansa voastră.

Alegeți o pietricică.

Urcați pe munte.

MULȚUMIRI

La fel ca inteligența artificială, această carte a fost scrisă pe parcursul mai multor ani. A început sub forma unor întrebări trimise prin mesaje text, a continuat sub forma unor conversații la masă și a escaladat la gradul de preocupare care mă urmărea la sală, la întâlniri sau în excursiile de weekend. O persoană-Brian Woolf- a acceptat această obsesie, mi-a permis să o cultiv și m-a sprijinit în toți acești ani. Brian a contribuit la cercetare, m-a ajutat să-mi exprim punctele de vedere și a stat noaptea târziu alături de mine pentru a-mi verifica fiecare pagină. Îi sunt profund recunoscătoare.

Cei Nouă Titani este produsul a sute de întâlniri tete-a-tete, interviuri și dineuri cu oameni care lucrează în inteligența artificială sau în domenii conexe

Sewell Chan, Noriyuki Shikata, Arfiya Eri, Joel Puckett, Erin McKean, Bill McBain, Frances Colon, Torfi Frans Olafsson, Latoya Peterson, Rob High, Anna Sekaran, Kris Schenck, Kara Snesko, Nadim Hossain, Megan Carroll, Elena Grewal, John Deutsch, Neha Narula, Toshi Ezoe, Masao Takahashi, Mary Madden, Shintaro Yamaguchi, Lorelei Kelly, Hiro Nozaki, Karen Ingram, Kirsten Gra-

ham, Francesca Rossi, Ben Johnson, Paola Antonelli, Yoav Schlesinger, Hardy Kagimoto, John Davidow, Rachel Sklar, Glynnis MacNicol, Yohei Sadoshima, și Eiko Ooka au fost suficient de generoși punându-mi la dispoziție timpul, perspectivele și descoperirile lor. Câțiva m-au recomandat altor persoane care lucrează în domeniul inteligenței artificiale și politicii, pentru a mă ajuta să înțeleg echilibrul geopolitic și oportunitățile și riscurile inteligenței artificiale.

L-am cunoscut pe Lt.Col Sea Thomas, maiorul în rezervă D.J. Skelton, directorul Defense Innovation Board Joshua Marcuse și analistul în securitate națională John Noonann prin intermediul US-Japan Leadership Foundation. Am petrecut multe zile împreună ca membri ai acestei societăți și sunt recunoscătoare fiecăruia dintre ei pentru răbdarea lor în a-mi explica viitorul războiului, rolul armatei americane în zona Pacificului și diferitele inițiative strategice ale Chinei. Sunt uimită în special de încercările lui Joshua de a reduce ecartul dintre Silicon Valley și Washington. El e unul dintre eroii din zilele noastre ai inteligenței artificiale.

Aspen Strategy Group mi-au oferit ocazia de a ține o prezentare despre viitorul inteligenței artificiale și geopolitică în cadrul întrunirii anuale de vară din Colorado, iar aceste conversații m-au ajutat să-mi conturez analiza. Le mulțumesc sincer lui Nicholas Burns, Condoleezza Rice, Joseph Nye și lui Jonathon Price pentru invitație, precum lui Carla Anne Robbins, Richard Danzig, James Baker, Wendy Sherman, Christian Brose, Eric Rosenbach, Susan Schwab, Ann-Marie Slaughter, Bob Zoellick, Philip Zelikow, Dov Zakheim, Laura Rosenberger și lui Mike Green pentru tot feedbackul lor prețios.

O bună parte din analiză s-a petrecut în campusul Stern School of Business a New York University, o casă deosebit de pri-

mitoare pentru studiul meu. Sunt recunoscătoare profesorului Sam Craig pentru că m-a acceptat în programul MBA și pentru că m-a sfătuit în ultimii ani. Nu pot să laud suficient studenții incredibil de inteligenți și creatori care au luat parte la cursurile mele. Trei dintre absolvenții de la Stern- Kriffy Perez, Elena Giralt și Roy Levkovitz- m-au ajutat în mod special să modelez viitorurile inteligenței artificiale.

Am avut șansa să am în anturajul meu un grup de înțelepți care să-mi ofere sfaturi. Toată munca mea este mai bună grație lor. Danny Stern mi-a schimbat viața acum câțiva ani când m-a invitat să-l cunosc în campusul NYU. El m-a învățat să gândesc mai exponențial și mi-a arătat cum să fac studiile mele accesibile unui public mult mai larg. Partenerul lui de la Stern Strategy Group, Mel Blake, a petrecut sute de ore învățându-mă, modelându-mi ideile și ajutându-mă să văd cu alți ochi lumea din jurul meu. Ei au reprezentat o sursă continuă de inspirație, motivație și (după cum știu) transpirație. James Geary și Ann Marie Lipinnski de la Harvard au fost incredibil de generoși vreme de mulți ani, permițându-mi să organizez întâlniri în care să discutăm despre viitor și să-mi îmbunătățesc metodologia de proiectare. James și Ann Marie sunt niște buni sfătuitori. Prietena mea dragă Maria Popova mă face să mă gândesc la lucruri mai mari, apoi contextualizează aceste idei prin cunoașterea ei enciclopedică a literaturii, artelor și științelor. Fata mea incredibilă, Petra Woolf, se întreabă tot timpul „ce ar fi dacă" aducându-mi aminte adesea de propriile prejudecăți cognitive legate de viitor. Sunt recunoscătoare, ca de obicei, profesorului Samuel Freedman de la Columbia University.

Mulțumiri profunde lui Cheryl Cooney, care lucrează neîncetat pentru mine și fără de care nu aș fi reușit mare lucru. Indife-

rent ce vor putea face IAG într-o bună zi, îmi e greu să-mi imaginez una care ar putea-o înlocui vreodată pe Cheryl. Emily Caufield-o femeie a cărei răbdare nu cunoaşte aparent nicio limită- e forța artistică din spatele tendințelor și scenariilor mele. Îi mulțumesc lui Phillip Blanchard pentru colaborarea în domeniul verificării datelor, redactării și compilării referințelor și notelor de subsol ale acestei cărți, și lui Mark Fortier, care s-a asigurat că atât cei care apar la ştiri cât și cei care le transmit au citit această carte, și ale cărui sfaturi s-au dovedit neprețuite în perioada lansării cărții.

În ultimul rând, datorez zettabaiți de recunoaştere lui Carol Franco, Kent Lineback și John Mahaney. Ca agent, Carol s-a ocupat de contractul pentru această carte. Ca prietenă însă, ea și cu soțul ei Kent m-au găzduit la casa lor frumoasă din Santa Fe pentru ca să putem dezvolta structura și tema centrală a celor Nouă Titani. Am petrecut zile și nopți condensând cercetarea și ideile mele în argumente cheie, iar între sesiunile de lucru ne plimbam prin oraş și aveam discuții interesante în restaurante grozave. Grație lui Carol l-am cunoscut pe editorul John Mahaney, cu care am avut privilegiul de a lucra la cartea mea anterioară. John e un editor ideal- pune multe întrebări, vrea informații de calitate și continuă să insiste până când analiza, exemplele și detaliile sunt numai bune. Am scris această carte pentru că voiam să modific discuția despre viitorul inteligenței artificiale, însă motivația mea nu era complet altruistă: a lucra din nou cu John a însemnat ocazia de a petrece un an cu el și de a-mi îmbunătăți stilul. John, Kent și Carol, sunteți o echipă formidabilă, și nu pot să cred cât de norocoasă sunt că v-am cunoscut.

BIBLIOGRAFIE

Abadi, M., A. Chu, I. Goodfellow, H. McMahan, I. Mironov, K. Talwar, and L. Zhang. "Deep Learning with Differential Privacy." In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS 2016)*, 308–318. New York: ACM Press, 2016. Abstract, ultima revizuire 24 octombrie 2016. https://arxiv.org/abs/1607.00133.

Ablon, L., and A. Bogart. *Zero Days, Thousands of Nights: The Life and Times of Zero-Day Vulnerabilities and Their Exploits*. Santa Monica, CA: RAND Corporation, 2017. https://www.rand.org/pubs/research_reports /RR1751.html.

Adams, S. S., et al. "Mapping the Landscape of Human-Level Artificial General Intelligence." *AI Magazine* 33, no. 1 (2012).

Agar, N. "Ray Kurzweil and Uploading: Just Say No!" *Journal of Evolution and Technology* 22 no. 1 (noiembrie 2011): 23–26. https://jetpress.org/v22 /agar.htm.

Allen, C., I. Smit, and W. Wallach. "Artificial Morality: Top-Down, Bottom-Up, and Hybrid Approaches." *Ethics and Information Technology* 7, no. 3 (2005).

Allen, C., G. Varner, and J. Zinser. "Prolegomena to Any Future Artificial Moral Agent." *Journal of Experimental and Theoretical Artificial Intelligence* 12, no. 3 (2000).

Allen, C., W. Wallach, and I. Smit. "Why Machine Ethics?" *IEEE Intelligent Systems* 21, no. 4 (2006).

Amdahl, G. M. "Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities." In *Proceedings of the AFIPS Spring Joint Computer Conference.* New York: ACM Press, 1967.

Anderson, M., S. L. Anderson, and C. Armen, eds. *Machine Ethics Technical Report FS-05-06.* Menlo Park, CA: AAAI Press, 2005.

Anderson, M., S. L. Anderson, and C. Armen. "An Approach to Computing Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006).

———. "MedE-thEx." In *Caring Machines Technical Report FS-05-02,* edited by T. Bickmore. Menlo Park, CA: AAAI Press, 2005.

———. "Towards Machine Ethics." In *Machine Ethics Technical Report FS-05- 06.* Menlo Park, CA: AAAI Press, 2005.

Anderson, S. L.„The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics." In *Machine Ethics.* Cambridge: Cambridge University Press, 2011.

Asimov, I. "Runaround." *Astounding Science Fiction* (martie 1942): 94–103.

Armstrong, S., A. Sandberg, and N. Bostrozz„Thinking Inside the Box." *Minds and Machines* 22, no. 4 (2012).

Axelrod, R. "The Evolution of Strategies in the Iterated Prisoner's Dilemma." In *Genetic Algorithms and Simulated Annealing,* edited by L. Davis. Los Altos, CA: Morgan Kaufmann, 1987.

Baars, B. J. "The Conscious Access Hypothesis." *Trends in Cognitive Sciences* 6, no. 1 (2002).

Babcock, J., et al. "Guidelines for Artificial Intelligence Containment." https:// arxiv.org/pdf/1707.08476.pdf.

Baier, C., and J. Katoen. *Principles of Model Checking.* Cambridge: MIT Press, 2008.

Bass, D. "AI Scientists Gather to Plot Doomsday Scenarios (and Solutions)." *Bloomberg,* 2 martie 2017. https://www.bloomberg.

com/news /articles /2017 -03-02/aiscientists-gather-to -plot-doomsday -scenarios -and -solutions.

Baum, S. D., B. Goertzel, and T. G. Goertzel. "How Long Until Human-Level AI? Results from an Expert Assessment." *Technological Forecasting and Social Change* 78 (2011).

Berg, P., D. Baltimore, H. W. Boyer, S. N. Cohen, R. W. Davis, D. S. Hogness, D. Nathans, R. Roblin, J. D. Watson, S. Weissman, and N. D. Zinder. "Potential Biohazards of Recombinant DNA Molecules." *Science* 185, no. 4148 (1974): 303.

Bostrom, N. "Ethical Issues in Advanced Artificial Intelligence." In *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence*, Vol. 2, edited by I. Smit and G. E. Lasker. Windsor, ON: International Institute for Advanced Studies in Systems Research and Cybernetics, 2003.

———. "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards." *Journal of Evolution and Technology* 9 (2002). http:// www.jetpress.org/volume9/risks.html.

———. "The Future of Human Evolution." In *Two Hundred Years After Kant, Fifty Years After Turing*, editat de C. Tandy, 339–371. Vol. 2 din *Death and Anti-Death*. Palo Alto, CA: Ria University Press, 2004.

———. "How Long Before Superintelligence?" *International Journal of Futures Studies*, nr. 2 (1998).

———. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

———. "The Superintelligent Will." *Minds and Machines* 22, no. 2 (2012).

———. "Technological Revolutions." In *Nanoscale*, editat de N. Cameron and M. E. Mitchell. Hoboken, NJ: Wiley, 2007.

Bostrom, N., and M. M. Ćirković, eds. *Global Catastrophic Risks*. New York: Oxford University Press, 2008.

Bostrom, N., and E. Yudkowsky. "The Ethics of Artificial Intelligence." In *Cambridge Handbook of Artificial Intelligence*, edited by K. Frankish and W. Ramsey. New York: Cambridge University Press, 2014.

Brooks, R. A. "I, Rodney Brooks, Am a Robot." *IEEE Spectrum* 45, no. 6 (2008).

Brundage, M., et al., "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation." https://arxiv.org/abs/1802.07228.

Brynjolfsson, E., şi A. McAfee. *The Second Machine Age*. New York: Norton, 2014.

Bryson, J., M. Diamantis, and T. Grant. "Of, For, and By the People: The Legal Lacuna of Synthetic Persons." *Artificial Intelligence and Law* 25, no. 3 (September 2017): 273–291.

Bueno de Mesquita, B., and A. Smith. *The Dictator's Handbook: Why Bad Behavior is Almost Always Good Politics*. New York: PublicAffairs, 2012.

Cassimatis N., E. T. Mueller, and P. H. Winston. "Achieving Human-Level Intelligence Through Integrated Systems and Research." *AI Magazine* 27, no. 2 (2006): 12–14. http://www.aaai.org/ojs/index.php/aimagazine/article /view/1876/1774.

Chalmers, D. J. *The Conscious Mind: In Search of a Fundamental Theory*. Philosophy of Mind Series. New York: Oxford University Press, 1996.

Chessen, M. *The MADCOM Future*. Washington, DC: Atlantic Council, 2017. http://www.atlanticcouncil.org/publications/reports/the-madcom -future.

Rapoartele Consiliului de Stat din China, disponibile pe platforma Consiliului de Stat al Republicii China www.gov.cn:

- Made in China 2025 (iulie 2015)
- State Council of a Next Generation Artificial Intelligence Development Plan (iulie 2017)

• Trial Working Rules on External Transfers of Intellectual Property Rights (martie 2018)
• Three-Year Action Plan on Blue Sky Days (iunie 2018)
• Three-Year Action Plan on Transportation Improvement (iunie 2018)
• State Council Approves Rongchang as National High-Tech Development Zone (martie 2018)
• State Council Approves Huainan as National High-Tech Development Zone (martie 2018)
• State Council Approves Maoming as National High-Tech Development Zone (martie 2018)
• State Council Approves Zhanjiang as National High-Tech Development Zone (martie 2018)
• State Council Approves Chuxiong as National High-Tech Development Zone (martie 2018)
• Three-Year Action Plan for Promoting Development of a New Generation Artificial Intelligence Industry 2018–2020 (decembrie 2017)
• Action Plan on the Belt Road Initiative (martie 2015)

Centre for New American Security. "Artificial Intelligence and Global Security Summit." https://www.cnas.org/events/artificial-intelligence-and-global -security-summit.

Core, M. G., et al. "Building Explainable Artificial Intelligence Systems." *AAAI* (2006): 1766–1773.

Crawford, K., and R. Calo. "There Is a Blind Spot in AI Research." *Nature*, 13 octombrie 2016. https://www.nature.com/news/there-is-a-blind-spot-in -ai -research-1.20805.

Dai, P., et al. "Artificial Intelligence for Artificial Artificial Intelligence." *AAAI Conference on Artificial Intelligence 2011*.

Dennett, D. C. "Cognitive Wheels." In *The Robot's Dilemma*, editat de Z. W. Pylyshyn. Norwood, NJ: Ablex, 1987.

Domingos, P. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.

Dvorsky, G. "Hackers Have Already Started to Weaponize Artificial Intelligence." *Gizmodo*, 2017. https://www.gizmodo.com.au/2017/09/hackers -have -already-started-toweaponize-artificial-intelligence/.

Dyson, G. *Darwin Among the Machines: The Evolution of Global Intelligence.* New York: Basic Books, 1997.

Eden, A., J. Søraker, J. H. Moor, and E. Steinhart, eds. *Singularity Hypotheses: A Scientific and Philosophical Assessment.* The Frontiers Collection. Berlin: Springer, 2012.

Evans, R., and J. Gao. "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%." DeepMind (blog), 20 iulie 2016. https://deepmind.com /blog/deepmind-ai-reducesgoogle-data-centre-cooling-bill-40/.

Fallows, J. *China Airborne.* New York: Pantheon, 2012.

Felten, E., şi T. Lyons. "The Administration's Report on the Future of Artificial Intelligence." Blog. 12 octombrie2016. https://obamawhitehouse .archives.gov/blog/2016/10/12/administrations-report -future-artificial -intelligence.

Floyd D. Spence National Defense Authorization Act for Fiscal Year 2001, Pub. L. No. 106–398, 114 Stat. 1654 (2001). http://www.gpo.gov/fdsys/pkg /PLAW-106publ398/html/PLAW-106publ398.htm.

French, H. *Midnight in Peking: How the Murder of a Young Englishwoman Haunted the Last Days of Old China.* Rev. ed. New York: Penguin Books, 2012.

Future of Life Institute. "Asilomar AI Principles." Text and signatories available online. https://futureoflife.org/ai-principles/.

Gaddis, J. L. *The Cold War: A New History.* New York: Penguin Press, 2006.

———. *On Grand Strategy.* New York: Penguin Press, 2018.

Gilder, G. F., and Ray Kurzweil. *Are We Spiritual Machines? Ray Kurzweil vs. the Critics of Strong AI.* edited by Jay Wesley Richards. Seattle:

Discovery Institute Press, 2001.

Goertzel, B., and C. Pennachin, eds. *Artificial General Intelligence.* Cognitive Technologies Series. Berlin: Springer, 2007. doi:10.1007/978-3-540-68677-4.

Gold, E. M. "Language Identification in the Limit." *Information and Control* 10, no. 5 (1967): 447–474.

Good, I. J. "Ethical Machines." *Intelligent Systems.* In vol. 10 of *Machine Intelligence*, edited by J. E. Hayes, D. Michie, and Y-H. Pao. Chichester, UK: Ellis Horwood, 1982.

———. "Speculations Concerning the First Ultraintelligent Machine." In vol. 6 of *Advances in Computers,* edited by F. L. Alt and M. Rubinoff. New York: Academic Press, 1965.

———. "Some Future Social Repercussions of Computers." *International Journal of Environmental Studies* 1, no. 1 (1970).

Greenberg, A. "The Jeep Hackers Are Back to Prove Car Hacking Can Get Much Worse." *Wired,* 1 August 2016. https://www.wired.com /2016/08 /jeep-hackers-return-high-speed-steering-acceleration-hacks/.

Harari, Y. N. *Homo Deus: A Brief History of Tomorrow.* New York: Harper, 2017.

Hilary, G. "The Professionalisation of Cyber Criminals." *INSEAD Knowledge* (blog), April 11, 2016. https://knowledge.insead.edu/blog/insead-blog/the -professionalisation-of-cyber-criminals-4626.

Hastie, T., R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics. New York: Springer, 2001.

Hofstadter, D. R. *Gödel, Escher, Bach: An Eternal Golden Braid.* New York: Basic Books, 1999.

Howard, P. K. *The Death of Common Sense: How Law Is Suffocating America.* New York: Random House, 1994.

Hua, Y. *China in Ten Words.* Tradus de A. H. Barr. New York: Pantheon Books, 2011.

Huang, W. *The Little Red Guard: A Family Memoir.* New York: Riverhead Books, 2012.

IEEE Spectrum. "Tech Luminaries Address Singularity." http://spectrum. ieee .org/computing/hardware/tech-luminaries-address-singularity.

IEEE Standards Association. "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems." https://standards.ieee.org/ develop/ind conn/ec/autonomous_systems.html.

Jo, YoungJu, et al. "Quantitative Phase Imaging and Artificial Intelligence: A Review." *Computing Research Repository* (2018). doi:abs/1806.03982.

Joy, B. "Why the Future Doesn't Need Us." *Wired,* April 1, 2000. http://www .wired.com/wired/archive/8.04/joy.html.

Kelly, K. *The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future.* New York: Viking, 2016.

Kirkpatrick, K. "Battling Algorithmic Bias." *Communications of the ACM* 59, no. 10 (2016): 16–17. https://cacm.acm.org/magazines/2016/10 /207759 -bat tling -algorithmic-bias/abstract.

Knight, W. "AI Fight Club Could Help Save Us from a Future of Super-Smart Cyberattacks." *MIT Technology Review*, July 20, 2017. https://www.tech nologyreview.com/s/608288/ai-fight-club-could-help-save-us-from-afu ture-of-supersmart-cyberattacks/.

———. "Response to Stephen Hawking." *Kurzweil Network*, September 5, 2001. http://www.kurzweilai.net/response-to-stephen-hawking.

———. *The Singularity Is Near.* New York: Viking, 2005.

Libicki, R. *Cyberspace in Peace and War.* Annapolis: Naval Institute Press, 2016.

Lin, J. Y. *Demystifying the Chinese Economy.* Cambridge, UK: Cambridge University Press, 2011.

Marcus, M. P., et al. "Building a Large Annotated Corpus of English: The Penn Treebank." *Computational Linguistics* 19, no. 2 (1993): 313–330.

Massaro, T. M., and H. Norton. "Siri-ously? Free Speech Rights and Artificial Intelligence." *Northwestern University Law Review* 110, no. 5 (2016): 1169–1194, Arizona Legal Studies Discussion Paper No. 15–29.

Minsky, M., P. Singh, and A. Sloman. "The St. Thomas Common Sense Symposium: Designing Architectures for Human-Level Intelligence." *AI Magazine* 25, no. 2 (2004).

Minsky, M. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind.* New York: Simon & Schuster, 2007.

———. *The Society of Mind.* New York: Simon & Schuster, 1985.

Neema, S. "Assured Autonomy." Defense Advanced Research Projects Agency. https://www.darpa.mil/program/assured-autonomy.

Osnos, E. *Age of Ambition: Chasing Fortune, Truth, and Faith in the New China.* New York: Farrar, Straus, and Giroux, 2015.

Petzold, C. *The Annotated Turing: A Guided Tour Through Alan Turing's Historic Paper on Computability and the Turing Machine.* Indianapolis, IN: Wiley Publishing, 2008.

Pylyshyn, Z. W., ed. *The Robot's Dilemma: The Frame Problem in Artificial Intelligence.* Norwood, NJ: Ablex, 1987.

Riedl, M. O. "The Lovelace 2.0 Test of Artificial Creativity and Intelligence." https://arxiv.org/pdf/1410.6142.pdf.

Schneier, B. "The Internet of Things Is Wildly Insecure—and Often Unpatchable." *Wired*, January 6, 2014. https://www.wired.com/2014/01/theres-no -good-way-to-patch-the-Internet-of-things-and-thats-a-huge-problem/.

Shannon, C., and W. Weaver. *The Mathematical Theory of Communica-*

tion. Urbana: University of Illinois Press, 1963.

Singer, P. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. London: Penguin Press, 2009.

Stanford University. "One Hundred Year Study on Artificial Intelligence (AI100)." https://ai100.stanford.edu/.

Toffler, A. *The Futurists*. New York: Random House, 1972.

Turing, A. M. "Intelligent Machinery, a Heretical Theory." Posthumous essay in *Philosophia Mathematica* 4, no. 3 (September 1, 1996): 256–260.

Tversky, A., and D. Kahneman. "The Framing of Decisions and the Psychology of Choice." *Science* 211, no. 4481 (1981).

Vinge, V. "The Coming Technological Singularity: How to Survive in the Post-Human Era." In *Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace*, NASA Conference Publication 10129 (1993): 11–22. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19940022855 _1994022855.pdf.

Wallach, W., and C. Allen. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press, 2009. doi:10.1093/acprof:oso /9780195374049.001.0001.

Weizenbaum, J. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W. H. Freeman, 1976.

Wiener, N. *The Human Use of Human Beings: Cybernetics and Society*. New York: Da Capo Press, 1950.

Yiwu, L. *The Corpse Walker: Real Life Stories, China from the Bottom Up*. Translated by W. Huang. New York: Anchor Books, 2009.

Yudkowsky, E. "AI as a Precise Art." Paper presented at the AGI Workshop 2006, Bethesda, MD, May 20, 2006.

NOTE

INTRODUCERE

1. Mozur, Paul. "Beijing Wants A.I. To Be Made In China By 2030," *New York Times*, 20 iulie 2017. https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html
2. Simonite, Tom. "Ex-Google Executive Opens A School For AI, With China's Help." *Wired*,5 aprilie 2018. https://www.wired.com/story/ex-google-executive-opens-a-school-for-ai-with-chinas-help/.
3. "Xinhua Headlines: Xi outlines blueprint to develop China's strength in cyberspace," *Xinhua*, April 213, 2018. http://www.xinhuanet.com /english/2018-04/21/c_137127374_2.htm.
4. Stephanie Nebehay, "U.N. says it has credible reports that China holds million Uighurs in secret camps," Reuters, 10 august 2018. https://www.reuters.com/article/us-china-rights-un/u-n-says-it-has-credible-reports-that-china-holds-million-uighurs-in-secret-camps -idUSKBN1KV1SU.
5. Mistreanu, Simina. "Life Inside China's Social Credit Laboratory," *Foreign Policy*, 3 aprilie, 2018. https://foreignpolicy.com/2018/04/03/life-inside-chinas-social-credit-laboratory/.
6. Idem.
7. "China Shames Jaywalkers Through Facial Recognition," *Phys.Org.*, 20 iunie 2017. https://phys.org/news/2017-06-china-shames-jaywalkers-facial-recognition.html.

CAPITOLUL 1: 1 MINTEA ŞI MAŞINILE: O FOARTE SCURTĂ ISTORIE A INTELIGENŢEI ARTIFICIALE

1. "The Seikilos Epitaph: The Oldest Song In The World". *Wired*, 29 octombrie 2009, https://www.wired.com/2009/10/the-seikilos-epitaph.
2. "Population Clock: World." *Census.Gov*., 2018, https://www.census.gov/popclock/world.
3. King, Elizabeth. "Clockwork Prayer: A Sixteenth-Century Mechanical Monk," *Virginia Commonwealth University*, Spring 2002, https://blackbird.vcu.edu/v1n1/nonfiction/king_e/prayer_introduction.htm.
4. Hobbes, Thomas. *De Corpore Politico, Or, The Elements Of Law Moral & Politick.*
5. Descartes, René. "Meditations on First Philosophy," Second Mediation §25, 1641. University of Connecticut. http://selfpace.uconn.edu/class/percep/DescartesMeditations.pdf.
6. Descartes, René. *Treatise of Man*, tr. T.S. Hall. Cambridge, Mass.: Harvard University Press, 1972.
7. Leibniz, Gottfried Wilhelm. *The Monadology*, tr. Robert Latta, 1898. https://www.plato-philosophy.org/wp-content/uploads/2016/07/The-Monadology-1714-by-Gottfried-Wilhelm-LEIBNIZ-1646-1716.pdf
8. Se consideră că prima folosire a cuvântului "computer" ar fi apărut într-o carte cu titlul *The Yong Mans Gleanings* scrisă de Richard Braithwaite în 1613. La acea vreme, computerele erau oamenii care efectuau calculele.
9. "Blaise Pascal." *Biography*. https://www.biography.com/people/blaise-pascal-9434176.
10. Leibniz scrie în *de Progressione Dyadica*: "Acest calcul [binar] ar putea fi implementat de o maşină...cu găuri care pot fi închise şi deschise. Ele vor fi deschise în locurile care corespund unui 1 şi vor rămâne închise în locurile care corespund unui 0. Prin porţile deschise mici cuburi sau pietricele pot cădea pe parcursuri, iar prin cele închise nu va putea trece nimic. Această [matrice a porţilor] se va muta de la o coloană la alta în funcţie de nevoie...!"
11. Leibniz scrie: "M-am gândit din nou la planul meu legat de un nou limbaj sau sistem de scriere pentru raţiune, care ar putea servi ca o unealtă de comunicare pentru toate naţiile. ...Dacă am dispune de un

asemenea instrument universal, am putea discuta problemele metafizice sau etice în acelaşi mod în care putem discuta matematica sau geometria. Acesta era obiectivul meu: orice neînţelegere nu ar trebui să fie altceva decât o eroare de calcul,...care să poată fi uşor corectată prin legile gramaticale ale acestui limbaj nou. Astfel, în cazul unei discuţii controversate, doi filosofi s-ar putea aşeza la masă şi ar putea calcula, spunând, ca doi matematicieni, ‚Hai să verificăm. ...”

12. “Apes to Androids: Is Man a Machine as La Mettrie Suggests?” http://www.charliemccarron.com/man_a_machine/.
13. Manabrea, Luigi. *Sketch of the Analytical Engine invented by Charles Babbage*. London: Richard and John E. Taylor, 1843.
14. Desmond MacHale, *The Life and Work of George Boole: A Prelude to the Digital Age*, New ed. (Cork University Press, 2014
15. Logicianul Martin Davis explică cel mai bine acest concept în cartea sa *The Universal Computer: The Road from Leibniz to Turing*: ”Turing ştia că un algoritm este de regulă precizat printr-o listă de reguli pe care o persoană o poate urma într-o manieră mecanică precisă, aidoma unei reţete dintr-o carte de bucate. El a putut demonstra că o astfel de persoană ar fi limitată la câteva acţiuni extrem de simple fără să modifice rezultatul calculului. Mai apoi, dovedind că nicio maşină care ar efectua doar aceste acţiuni elementare nu ar putea să stabilească dacă o concluzie dată rezultă din premisele date... el a putut concluziona că nu există niciun algoritm pentru această Entscheidungsproblem.”
16. Davis, Martin. *The Universal Computer: The Road from Leibniz to Turing, Third Edition.*” Boca Raton, Fla: CRC Press, 2018.
17. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” Intelligence,” 31 august 1955, Stanford University, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.
18. În cadrul propunerii pentru proiectul de cercetare de vară privind inteligenţa artificială de la Dartmouth, McCarthy, Minsky, Rochester şi Shannon au propus ca următorii oameni să fie invitaţi. Am redat aici lista originală publicată în 1955, care include şi numele companiilor şi adresele. Nu toţi invitaţii au putut participa.

Adelson, Marvin
Hughes Aircraft Company
Airport Station, Los Angeles, CA

Ashby, W. R.
Barnwood House
Gloucester, England

Backus, John
IBM Corporation
590 Madison Avenue
New York, NY

Bernstein, Alex
IBM Corporation
590 Madison Avenue
New York, NY

Bigelow, J. H.
Institute for Advanced Studies
Princeton, NJ

Elias, Peter
R. L. E., MIT
Cambridge, MA

Duda, W. L.
IBM Research Laboratory
Poughkeepsie, NY

Davies, Paul M.
1317 C. 18th Street
Los Angeles, CA

Fano, R. M.
R. L. E., MIT
Cambridge, MA

Farley, B. G.
324 Park Avenue
Arlington, MA

Galanter, E. H.
University of Pennsylvania
Philadelphia, PA

Gelernter, Herbert
IBM Research
Poughkeepsie, NY

Glashow, Harvey A.
1102 Olivia Street
Ann Arbor, MI

Goertzal, Herbert
330 West 11th Street
New York, New York

Hagelbarger, D.
Bell Telephone Laboratories
Murray Hill, NJ

Miller, George A.
Memorial Hall
Harvard University
Cambridge, MA

Harmon, Leon D.
Bell Telephone Laboratories
Murray Hill, NJ

Holland, John H.
E. R. I.
University of Michigan
Ann Arbor, MI

Holt, Anatol
7358 Rural Lane
Philadelphia, PA

Kautz, William H.
Stanford Research Institute
Menlo Park, CA

Luce, R. D.
427 West 117th Street
New York, NY

MacKay, Donald
Department of Physics
University of London
London, WC2, England

McCarthy, John
Dartmouth College
Hanover, NH

McCulloch, Warren S.
R.L.E., M.I.T.
Cambridge, MA

Melzak, Z. A.
Mathematics Department
University of Michigan
Ann Arbor, MI

Minsky, M. L.
112 Newbury Street
Boston, MA

More, Trenchard
Department of Electrical Engineering
MIT
Cambridge, MA

Nash, John
Institute for Advanced Studies
Princeton, NJ

Newell, Allen
Department of Industrial Administration
Carnegie Institute of Technology

Pittsburgh, PA

Robinson, Abraham
Department of Mathematics
University of Toronto
Toronto, Ontario, Canada

Rochester, Nathaniel
Engineering Research Laboratory
IBM Corporation
Poughkeepsie, NY

Rogers, Hartley, Jr.
Department of Mathematics
MIT
Cambridge, MA

Rosenblith, Walter
R.L.E., M.I.T.
Cambridge, MA

Rothstein, Jerome
21 East Bergen Place
Red Bank, NJ

Sayre, David
IBM Corporation
590 Madison Avenue
New York, NY

Schorr-Kon, J.J.
C-380 Lincoln Laboratory, MIT
Lexington, MA

Shapley, L.
Rand Corporation
1700 Main Street
Santa Monica, CA

Schutzenberger, M.P.
R.L.E., M.I.T.
Cambridge, MA

Selfridge, O. G.
Lincoln Laboratory, M.I.T.
Lexington, MA

Shannon, C. E.
R.L.E., M.I.T.
Cambridge, MA

Shapiro, Norman
Rand Corporation
1700 Main Street
Santa Monica, CA

Simon, Herbert A.
Department of Industrial Administration
Carnegie Institute of Technology
Pittsburgh, PA

Solomonoff, Raymond J.
Technical Research Group
17 Union Square West
New York, NY

Steele, J. E., Capt. USAF
Area B., Box 8698
Wright-Patterson AFB
Ohio

Webster, Frederick
62 Coolidge Avenue
Cambridge, MA

Moore, E. F.
Bell Telephone Laboratory
Murray Hill, NJ

Kemeny, John G.
Dartmouth College
Hanover, NH

19. Am compus o listă foarte scurtă de femei şi oameni de culoare care ar fi putut contribui substanţial la atelierul de la Dartmouth, dar au fost omişi. Această listă nu este exhaustivă. Aş fi putut continua pe zeci de pagini. Ea este emblematică pentru oamenii inteligenţi, capabili şi creativi care nu au fost invitaţi.

James Andrews, matematician şi profesor la Florida State University, specializat în teoria grupurilor şi teoria nodurilor.

Jean Bartik, matematiciană, unul dintre programatorii computerului ENIAC.

Albert Turner Bharucha-Reid, matematician şi teoretician cu contribuţii importante în domeniul lanţurilor Markov, teoriei probabilităţii şi statisticii.

David Blackwell, statistician şi matematician cu contribuţii importante în teoria jocului, teoria informaţiei, teoria probabilităţii şi statistica bayesiană.

Mamie Phipps Clark, doctor şi psiholog social, cu experienţă în domeniul conştiinţei de sine.

Thelma Estrin, pionier al aplicării sistemelor informatice în cercetarea neurofiziologică şi cerebrală. Cercetătoare în departamentul de electroencefalografie al Institutului de Neurologie de la Columbia Presbyterian în perioada proiectului de cercetare de vară de la Dartmouth.

Evelynn Boyd Granville, doctor în matematică, a elaborat programele folosite pentru analiza traiectoriilor în cadrul primelor misiuni cu astronauţi trimise în spaţiu şi spre lună.

Betty Holberton, matematiciană, unul dintre programatorii computerului ENIAC. Ea a inventat punctele de oprire în depanarea programelor.

Grace Hopper, informaticiană, creatoare a limbajului COBOL, folosit până în zilele noastre.

Mary Jackson, ingineră și matematiciană, prima femeie inginer de culoare de la NASA.

Kathleen McNulty, matematiciană, unul dintre programatorii computerului ENIAC.

Marlyn Meltzer, matematiciană, unul dintre programatorii computerului ENIAC, primul computer programabil integral electronic.

Rozsa Peter, matematiciană, fondatoare a teoriei funcțiilor recursive.

Frances Spence, matematiciană, unul dintre programatorii computerului ENIAC.

Ruth Teitelbaum, matematiciană, unul dintre programatorii computerului ENIAC. A calculat împreună cu Marlyn Meltzer ecuații pentru traiectorii balistice.

Dorothy Vaughan, matematiciană, supervizoare în 1949 a grupului West Area Computers.

Jesse Ernest Wilkins Jr., fizician nuclear, inginer mecanic și matematician, a devenit cel mai tânăr student al universității din Chicago la vârsta de 13 ani.

20. "The Dartmouth Workshop --as planned and as it happened," http://www-formal.stanford.edu/jmc/slides/dartmouth/dartmouth/node1.html.
21. "The Dartmouth AI Archives," http://raysolomonoff.com/dartmouth/.
22. Irving John Good, "Speculations Concerning the First Ultraintelligent Machine," *Advances in Computers,* Volume 6 (1966): 31–88, https://www.sciencedirect.com/science/article/pii/S0065245808604180?via%3Dihub.
23. Weizenbaum, Joseph, "ELIZA — A Computer Program for the Study of Natural Language Communication Between Man and Machine," *Computational Linguistics,* January, 1966, http://web.stanford.edu/class/cs124/p36-weizenabaum.pdf
24. Stenograma completă poate fi accesată pe Github: https://github.com/codeanticode/eliza
25. "The Chicago Tribune Archives," *Chicago Tribune,* 18 iunie 1963, http://www.chicagotribune.com/tribune-archives/.
26. H.A. Simon and A. Newell, "Heuristic Problem Solving: The Next

Advance in Operations Research," *Operations Research 6* (1958): 1- 10.

27. McCarthy însuşi a dorit să lucreze cu grupul în ce priveşte ideile proprii pentru reprezentarea cunoştinţelor de bun simţ şi raţiunea, dar odată ce grupul a fost asamblat, şi-a dat seama că din grup lipseau anumiţi gânditori cheie. (În cazul lui, se aştepta să fi existat şi logicieni.)
28. "2001 Is Near, but 'Hal' Is Not," *Los Angeles Times, 7 ianuarie 1977*, http://articles.latimes.com/1997-01-07/news/mn-16139_1_artificial-intelligence
29. "Language and Machines: Computers in Translation and Linguistics," The National Academies of Science, Engineering and Medicine, 1966, https://www.nap.edu/read/9547/chapter/1
30. Lighthill, James. "Artificial Intelligence: A General Survey," *Informatics*, iulie 1972, http://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm.
31. "Thinking Allowed, Conversations On the Leading Edge of Knowledge and Discovery, with Dr. Jeffrey Mishlove," *Intuition.org*, 1998, http://www.intuition.org/txt/minsky.htm.
32. Idem.
33. Iarna inteligenţei artificiale, care a durat trei decenii, a inclus noi previziuni- de data aceasta sub forma unor avertismente- pentru viitor. În cartea sa *Computer Power and Human Reason*, Weizenbaum susţinea că deşi inteligenţa artificială este posibilă, nu ar trebui să permitem niciodată calculatoarelor să ia decizii importante pentru că lor le vor lipsi întotdeauna calităţi umane precum compasiunea sau inteligenţa. Weizenbaum face aici o distincţie esenţială între a decide şi a alege. Decizia este o activitate legată de calcul, care poate fi programată. Alegerea este însă produsul judecăţii, nu al calculului. Capacitatea de a alege este ceea ce ne conferă în ultimă instanţă umanitatea. Filosoful John Searle de la University of California, Berkeley susţine în articolul său "Minds, Brains and Programs" că o inteligenţă artificială generală, sau, în termenii lui, "puternică" este implauzibilă. Searle spunea că un program nu poate conferi unui computer o "minte", o "înţelegere" sau o "conştiinţă", indiferent cât de asemănător cu oamenii s-ar putea comporta acel program.

34. Schaeffer, Jonathan, Robert Lake, Paul Lu, Martin Bryant. "CHINOOK: The World Man-Machine Checkers Champion," *AI Magazine*, primăvara, 1966, https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1208/1109
35. Goldfarb, Ari, Daniel Trefler. "AI and International Trade," *The National Bureau of Economic Research*, January, 2018, http://www.nber.org/papers/w24254.pdf.
36. Toby Manning, "AlphaGo," *British Go Journal* 174 (Winter 2015–2016): 15, https://www.britgo.org/files/2016/deepmind/BGJ174-AlphaGo.pdf.
37. Byford, Sam. "AlphaGo retires from competitive Go after defeating world number one 3-0," *The Verge*, May 27, 2017, https://www.theverge.com/2017/5/27/15704088/alphago-ke-jie-game-3-result-retires-future.
38. Silver, David, Julian Schrittwieser, Karen Simonyan, et al. "Mastering the Game of Go without Human Knowledge," DeepMind, https://deepmind.com/documents/119/agz_unformatted_nature.pdf
39. Idem.
40. Ibidem.
41. Afirmația a fost făcută de programatorul șef al lui Zero, David Silver, în cadrul unei conferințe de presă.
42. Byford, Sam. "AlphaGo retires from competitive Go after defeating world number one 3-0," *The Verge*, May 27, 2017, https://www.theverge.com/2017/5/27/15704088/alphago-ke-jie-game-3-result-retires-future.
43. Novet, Jordan. "Google is finding ways to make money from Alphabet's DeepMind A.I. technology," *CNBC*, 31 martie 2018, https://www.cnbc.com/2018/03/31/how-google-makes-money-from-alphabets-deepmind-ai-research-group.html
44. Cerejo, Roydon. "Google Duplex: Understanding the Core Technology Behind Assistant's Phone Calls," *Gadgets 360*, 10 mai 2018, https://gadgets.ndtv.com/apps/features/google-duplex-google-io-ai-google-assistant-1850326
45. Quoc Le, Barret Zoph."Using Machine Learning to Explore Neural Network Architecture, *Google AI Blog*, 17 mai 2017, https://ai.google-

blog.com/2017/05/using-machine-learning-to-explore.html.

46. Schema Winograd, propusă de informaticianul canadian Hector Levesque în 2011, prezintă o alternativă la testul Turing pentru a măsura capacitățile unei inteligențe artificiale. Schema, pe care Levesque și colegii au dennumit-o în onoarea informaticianului Terry Winograd de la Stanford se axează pe raționamentele de bun simț, provocând inteligența artificială să răspundă la o întrebare cu pronume ambigue. Accentul pus pe depășirea oamenilor în competiție directă a condus la neglijarea altor metode de a măsura și a face inteligența artificială să progreseze. Schema Winograd a fost gândită ca un test multidimensional pentru că pentru a-ș putea trece e nevoie de mai mult decât de un set vast de date. Ernest Davis, Leora Morgenstern și Charles Ortiz, trei informaticieni de la NYU, au propus o provocare a schemei Winograd, care are loc o dată pe an. Pe pagină facultății există un exemplu excelent:
https://cs.nyu.edu/faculty/davise/papers/WinogradSchemas/WS.htm. Last accessed September 5, 2018.

Consilierii municipali au refuzat să acorde protestatarilor autorizația de a demonstra pentru că aceștia [se temeau de/susțineau] violențe. Dacă cuvântul e "se temeau" atunci "aceștia" se referă la consilieri. Dacă e "susțineau" atunci aceștia se referă la protestatari.

În cadrul articolului, LEvesque afirmă că schemele Winograd ar trebui să respecte următoarele constrângeri:

- Să poată fi ușor dezlegate de un cititor uman (în mod ideal, atât de ușor încât cititorul să nici nu sesizeze prezența vreunei ambiguități.
- Să nu poată fi rezolvate prin tehnici simple cum ar fi restricțiile de selecție
- Să fie imune la Google; cu alte cuvinte, nu există un test evident statistic aplicat bazelor de date cu text care să poată dezlega frazele corect

CAPITOLUL 2: LUMEA INSULARĂ A TRIBURILOR INTELIGENȚEI ARTIFICIALE

1. Mike Isaac şi Sheera Frenkel, "Facebook Security Breach Exposes Accounts of 50 Million Users," *New York Times*, 28 septembrie 2018, https://www.nytimes.com/2018/09/28/technology/facebook-hack-data -breach.html.
2. Casey Newton, "Facebook Portal's Claims to Protect User Privacy Are Falling Apart, *The Verge,* 17 octombrie 2018, https://www .theverge.com/2018/10/17/17986992/facebook-portal-privacy - claims-ad-targeting.
3. "AMA: We are the Google Brain team. We'd love to answer your questions about machine learning," *Reddit,* 2016, https://www.reddit.com/r/MachineLearning/comments/4w6tsv/ama_ we_are_the_google_brain_team_wed_love_to/.
4. Idem.
5. "Diversity," *Google,* https://diversity.google/.
6. Tiku, Nitasha. "Google's Diversity Stats Are Still Very Dismal," Wired,14 august 2018, https://www.wired.com/story/googles-employee-diversity-numbers-havent-really-improved/.
7. Daisuke Wakabayashi and Katie Benner, "How Google Protected Andy Rubin, the 'Father of Android,'" *New York Times*, 25 octombrie 2018, https://www.nytimes.com/2018/10/25/technology/google-sexual -harassment-andy-rubin.html.
8. Broockman, David, Greg F. Ferenstein, Neil Malhotra. "The Political Behavior of Wealthy Americans: Evidence from Technology Entrepreneurs," *Stanford Graduate School of Business*, 9 decembrie 2017, https://www.gsb.stanford.edu/faculty-research/working-papers/ political-behavior-wealthy-americans-evidence-technology.
9. "ICYMI: RNC Chairwoman & Brad Parscale Demand Answers from Facebook & Twitter," *Republican National Committee*, 24 mai 2018, https://www.gop.com/icymi-rnc-chairwoman-brad-parscale-demand-answers-from-facebook-twitter.
10. Conger, Kate, Sheera Frenkel. "Dozens at Facebook Unite to Challenge Its 'Intolerant' Liberal Culture," *New York Times*, 28 august 2018, https://www.nytimes.com/2018/08/28/technology/inside-face-

book-employees-political-bias.html.
11. Rocha, Veronica. "Crime-fighting robot hits, rolls over child at Silicon Valley mall," *Los Angeles Times*, 14 iulie 2016, http://www.latimes.com/local/lanow/la-me-ln-crimefighting-robot-hurts-child-bay-area-20160713-snap-story.html.
12. Benson, Julian. "Elite's AI Created Super Weapons and Started Hunting Players. Skynet is Here," *Kotaku*, 3 iunie 2016, http://www.kotaku.co.uk/2016/06/03/elites-ai-created-super-weapons-and-started-hunting-players-skynet-is-here.
13. Bucks County Courier Times Newspaper Archives, 20 august 1974, p. 9
14. Există un articol grozav despre Pryor înainte să devină celebru. McPherson, James. "The new comic style of Richard Pryor," *New York Times*, April 27, 1975
15. Vance, Ashlee. "How We Got Here," *Bloomberg Businessweek, 21* mai 2018, https://www.scribd.com/article/379513106/How-We-Got-Here.
16. "Computer Science," *Stanford University*, 2018-2019, https://explore-degrees.stanford.edu/schoolofengineering/computerscience/#bachelortext.
17. "Vector Representations of Words," *TensorFlow*, https://www.tensorflow.org/tutorials/representation/word2vec.
18. Bolukbasi, Tolga, Kai-Wei Chang, et al.,"Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings," *Cornell University Library*, 21 iulie 2016, https://arxiv.org/abs/1607.06520.
19. Saltiel, Natalie. "The Ethics and Governance of Artificial Intelligence," *MIT Media Lab*, 16 noiembrie 2017, https://www.media.mit.edu/courses/the-ethics-and-governance-of-artificial-intelligence/.
20. Puteți urmări cursurile aici: https://www.media.mit.edu/courses/the-ethics-and-governance-of-artificial-intelligence/.
21. Ashcraft, Catherine, Brad McLain, Elizabeth Eger. "Women in Tech: The Facts," *National Center for Women & Information Technology*, 2016, https://www.ncwit.org/sites/default/files/resources/womenintech_facts_fullreport_05132016.pdf.
22. "Degrees in computer and information sciences conferred by degree-granting institutions, by level of degree and sex of student:

1970-71 through 2010-11," *National Center for Education Statistics*, https://nces.ed.gov/programs/digest/d12/tables/dt12_349.asp.
23. Doctor's degrees conferred by postsecondary institutions, by race/ethnicity and field of study: 2013-14 and 2014-15," *National Center for Education Statistics*, https://nces.ed.gov/programs/digest/d16/tables/dt16_324.25.asp?current=yes.
24. Mims, Christopher. "What the Google Controversy Misses: The Business Case for Diversity," *Wall Street Journal*, 13 august 2017, https://www.wsj.com/articles/what-the-google-controversy-misses-the-business-case-for-diversity-1502625603.
25. Hempel, Jessi. "Melinda Gates and Fei-Fei Li Want to Liberate AI from "Guys With Hoodies," *Wired*, 4 mai 2017, https://www.wired.com/2017/05/melinda-gates-and-fei-fei-li-want-to-liberate-ai-from-guys-with-hoodies/.
26. Meng Jing. "China looks to school kids to win the global AI race," *South China Morning Post, International Edition*, 3 mai 2018, https://www.scmp.com/tech/china-tech/article/2144396/china-looks-school-kids-win-global-ai-race.
27. "China launches first university program to train intl AI talents," *Zhongguancun Science Park*, 4 aprilie 2018, http://www.chinadaily.com.cn/m/beijing/zhongguancun/2018-04/04/content_35979394.htm.
28. Barboza, David. "The Rise of Baidu (That's Chinese for Google)," *New York Times*, 1 septembrie 2006, https://www.nytimes.com/2006/09/17/business/yourmoney/17baidu.html.
29. "Rise Of China's Big Tech In AI: What Baidu, Alibaba, And Tencent Are Working On," *CB Insights*, 26 aprilie 2018, https://www.cbinsights.com/research/china-baidu-alibaba-tencent-artificial-intelligence-dominance/.
30. Lucas, Louise. "The Chinese Communist party entangles big tech," *Financial Times*, 18 iulie 2018, https://www.ft.com/content/5d0af3c4-846c-11e8-a29d-73e3d454535d.
31. "A Hong Kong Newspaper on a Mission to Promote China's Soft

Power," *New York Times*, 31 martie 2018, https://www.nytimes.com/2018/03/31/world/asia/south-china-morning-post-hong-kong-alibaba.html.

32. Farhi, Paul."Washington Post closes sale to Amazon founder Jeff Bezos," *Washington Post*, 1 octombrie 2013, https://www.washington-post.com/business/economy/washington-post-closes-sale-to-amazon-founder-jeff-bezos/2013/10/01/fca3b16a-2acf-11e3-97a3-ff2758228523_story.html?noredirect=on&utm_term=.3d04830eab75
33. Lim, Jason. "WeChat Is Being Trialled To Make Hospitals More Efficient In China," *Forbes*,16 iunie 2014 https://www.forbes.com/sites/jlim/2014/06/16/wechat-is-being-trialed-to-make-hospitals-more-efficient-in-china/#63a2dd3155e2
34. "Rise Of China's Big Tech In AI: What Baidu, Alibaba, And Tencent Are Working On," *CB Insights, 26 a*prilie 2018, https://www.cbinsights.com/research/china-baidu-alibaba-tencent-artificial-intelligence-dominance/.
35. Kharpal, Arjun. "China's Tencent surpasses Facebook in valuation a day after breaking $500 billion barrier," *CNBC*, 21 noiembrie 2017, https://www.cnbc.com/2017/11/21/tencent-surpasses-facebook-in-valuation.html.
36. Rutherford, Sam. "5 Things to Know About Tencent, the Chinese Internet Giant That's Worth More Than Facebook Now," *Gizmodo*, November 27, 2017, https://gizmodo.com/5-things-to-know-about-tencent-the-chinese-internet-gi-1820767339.
37. Fannin, Rebecca. "China Releases a Tech Dragon: the BAT," *Techonomy*, 23 mai 2018, https://techonomy.com/2018/05/china-releases-tech-dragon-bat/.
38. "Mobile Fact Sheet," *Pew Research Center*, 5 februarie 2018, http://www.pewinternet.org/fact-sheet/mobile/.
39. Yurieff, Kaya. "Amazon's Cyber Monday was its biggest sales day ever," *CNN Money*, 29 noiembrie 2017, https://money.cnn.com/2017/11/29/technology/amazon-cyber-monday/index.html.
40. Wang, Helen H. "Alibaba's Singles' Day By The Numbers: A Record $25 Billion Haul," *Forbes*, 12 noiembrie 2017.

41. Fannin, Rebecca. "China Releases a Tech Dragon: the BAT," *Techonomy*, 23 mai 2018, https://techonomy.com/2018/05/china-releases-tech-dragon-bat/.
42. Brown, Michael, Pavneet Singh. "China's Technology Transfer Strategy," *Defense Innovation Unit Experimental*, februarie 2017, https://new.reorg-research.com/data/documents/20170928/59ccf7de70c2f.pdf.
43. Pentru cei care doresc să acceseze textul integral al celui de-al 13-lea plan cincinal, el poate fi consultat la adresa http://www.gov.cn/xinwen/2016-03/17 /content_5054992.htm. 17 martie 2016, traducere în limba engleză.
44. J.P. "What is China's belt and road initiative?," *The Economist*, 15 mai 2017, https://www.economist.com/the-economist-explains/2017/05/14/what-is-chinas-belt-and-road-initiative.
45. Babones, Salvatore. "China's Middle Class Is Pulling Up the Ladder Behind Itself," *Foreign Policy*, 1 februarie 2018, https://foreignpolicy.com/2018/02/01/chinas-middle-class-is-pulling-up-the-ladder-behind-itself/.
46. "The American Middle Class Is Losing Ground," *Pew Research Center*, 9 decembrie 2015, http://www.pewsocialtrends.org/2015/12/09/the-american-middle-class-is-losing-ground/.
47. Martin, Emmie. "70% of Americans consider themselves middle class—but only 50% are," *CNBC*, 30 iunie 2017, https://www.cnbc.com/2017/06/30/70-percent-of-americans-consider-themselves-middle-class-but-only-50-percent-are.html
48. Bhattarai, Abha. "China asked Marriott to shut down its website. The company complied," *Washington Post*, 18 ianuarie 2018.
49. Jacobson, Louis. "Yes, Donald Trump did call climate change a Chinese hoax," *PolitiFact*, 3 iunie 2016, https://www.politifact.com/truth-o-meter/statements/2016/jun/03/hillary-clinton/yes-donald-trump-did-call-climate-change-chinese-h/.
50. Greenstone, Michael. "Four Years After Declaring War on Pollution, China Is Winning," *New York Times*, 12 martie 2018, https://www.nytimes.com/2018/03/12/upshot/china-pollution-environment-longer-lives.html

https://www.nytimes.com/2018/03/12/upshot/china-pollution-environment-longer-lives.html.
51. Fordham, Carl Gene. "20 Actually Useful Chengyu," *Carlgene.com*, 14 August 2008, http://carlgene.com/blog/2010/07/20-actually-useful-chengyu-%E6%88%90%E8%AF%AD.
52. Chen, Stephen. "China takes surveillance to new heights with flock of robotic Doves, but do they come in peace?" *South China Morning Post*, 24 iunie 2018, https://www.scmp.com/news/china/society/article/2152027/china-takes-surveillance-new-heights-flock-robotic-doves-do-they
53. Stewart, Phil. "China racing for AI military edge over U.S.: report," *Reuters*, 27 noiembrie 2017, https://www.reuters.com/article/us-usa-china-ai/china-racing-for-ai-military-edge-over-u-s-report-idUSKBN1DS0G5.
54. Conger, Kate. "Google Employees Resign in Protest Against Pentagon Contract," *Gizmodo*, 14 mai 2018, https://gizmodo.com/google-employees-resign-in-protest-against-pentagon-con-1825729300.
55. Nitasha Tiku, "Amazon's Jeff Bezos Says Tech Companies Should Work with the Pentagon," *Wired*, October 15, 2018. https://www.wired.com /story/amazons-jeff-bezos-says-tech-companies-should-work-with-the -pentagon/.
56. Stewart, Phil. "China racing for AI military edge over U.S.: report," *Reuters*, 27 noiembrie 2017, https://www.reuters.com/article/us-usa-china-ai/china-racing-for-ai-military-edge-over-u-s-report-idUSKBN1DS0G5.
7. "China issues guideline on artificial intelligence development," *The State Council, People's Republic of China*, 20 iulie 2017, http://english.gov.cn/policies/latest_releases/2017/07/20/content_281475742458322.htm.
58. "Key AI guidelines unveiled," *The State Council, People's Republic of China*, 15 decembrie 2017, http://english.gov.cn/state_council/ministries/2017/12/15/content_281475977265006.htm.
59. Kania, Elsa B. "China's AI Giants Can't Say No to the Party," *Foreign Policy*, 2 august 2018, https://foreignpolicy.com/2018/08/02/chinas-ai-giants-cant-say-no-to-the-party/.

60. Idem.
61. Ibidem.
62. Pomfret, John. "China's new surveillance state puts Facebook's privacy problems in the shade," *Washington Post*, 27 martie 2018, https://www.washingtonpost.com/news/global-opinions/wp/2018/03/27/chinas-new-surveillance-state-puts-face-books-privacy-problems-in-the-shade/?noredirect=on&utm_term=.d8cc0acd08b1.
63. Wright, Nicholas. "How Artificial Intelligence Will Reshape the Global Order," *Foreign Affairs*, 10 iulie 2018, https://www.foreignaffairs.com/articles/world/2018-07-10/how-artificial-intelligence-will-reshape-global-order.
64. Zhang Hongpei. "Many netizens take issue with Baidu CEO's comments on data privacy," *Global Times*, 26 martie 2018, http://www.globaltimes.cn/content/1095288.shtml
65. Zhong, Raymond. "Chinese Tech Giant on Brink of Collapse in New U.S. Cold War," *New York Times*, 9 mai 2018
66. Sacks, Samm. "Beijing Wants to Rewrite the Rules of the Internet," *The Atlantic*, 19 iunie 2018, https://www.theatlantic.com/international/archive/2018/06/zte-huawei-china-trump-trade-cyber/563033/.
67. Idem.
68. Ibidem.
69. "The Recruitment Program for Innovative Talents (Long Term)," *The Thousand Talents Plan*, http://1000plan.org/en/.
70. Marquand, Michael. "The Trump Administration Plays Catch-Up on Artificial Intelligence," *Wired*, 11 mai 2018, https://www.wired.com/story/trump-administration-plays-catch-up-artificial-intelligence/.
71. Levy, Ari. "Dropbox is going public: Here's who's making money," *CNBC*, 23 februarie 2018, https://www.cnbc.com/2018/02/23/dropbox-is-going-public-heres-whos-making-money.html.
72. Gramlich, John. "5 facts about Americans and Facebook," *Pew Research Center*, 10 aprilie 2018, http://www.pewresearch.org/fact-tank/2018/04/10/5-facts-about-americans-and-facebook/
73. Weise, Elizabeth. "Amazon Prime is popular, but in three-quarters of

all U.S homes? That's open to debate," *USA Today*, 20 octombrie 2017, https://www.usatoday.com/story/tech/2017/10/20/amazon-prime-big-though-how-big-no-one-knows/784695001/.
74. "Mobile Fact Sheet," *Pew Research Center*, 5 februarie 2018, http://www.pewinternet.org/fact-sheet/mobile/.
75. *Tensorflow*, https://github.com/tensorflow/tensorflow.
76. "Microsoft to acquire GitHub for $7.5 billion," *Microsoft News Center*, 4 iunie 2018, https://news.microsoft.com/2018/06/04/microsoft-to-acquire-github-for-7-5-billion/.
77. Novet, Jordan. "Why tech companies are racing each other to make their own custom A.I. chips," *CNBC*, 21 aprilie 2018, https://www.cnbc.com/2018/04/21/alibaba-joins-google-others-in-making-custom-ai-chips.html.
78. The full paper can be accessed at https://graphics.axios.com/pdf / PlatformPolicyPaper.pdf?_ga=2.167458877.2075880604.1541172609 -1964512884.1536872317.
79. Aceste mesaje pot fi accesate la adresa https://twitter.com/tim_cook/status /1055035534769340418.

CAPITOLUL 3: O MIE DE TĂIETURI CU HÂRTIA: CONSECINȚELE NEINTENȚIONATE ALE INTELIGENȚEI ARTIFICIALE

1. "An Owners' Manual for Google Shareholders," *Alphabet Investor Relations*, https://abc.xyz/investor/founders-letters/2004/ipo-letter.html.
2. Idem.
3. "Leadership Principles," *Amazon Jobs*, https://www.amazon.jobs/principles.
4. "Focus on Impact," Facebook, 8 septembrie 2015, https://www.facebook.com/facebookcareers/photos/a.1655178611435493.1073741828.1633466236940064/1655179928102028/?type=3&theater.
5. "Core Values," Tencent, https://www.tencent.com/en-us/culture.html.
6. "Culture and Values," Alibaba Group, https://www.alibabagroup.com/en/about/culture.
7. Bergen, Mark. "Google Engineers Refused to Build Security Tool to Win Military Contracts," *Bloomberg*, 21 iunie 2018, https://www.bloomberg.com/news/articles/2018-06-21/google-

engineers-refused-to-build-security-tool-to-win-military-contracts.
8. Pichai, Sundar. "AI at Google: our principles," *Google*, 7 iunie 2018, https://www.blog.google/technology/ai/ai-principles/.
9. "Quick Facts, *U.S. Census Bureau*, 1 iulie 2017, https://www.census.gov/quickfacts/fact/table/US/PST045217.
10. McCormack, Alan, John Rusnak, Carliss Baldwin. "Exploring the Duality between Product and Organizational Architectures: A Test of the "Mirroring" Hypothesis," *Harvard Business School*, https://www.hbs.edu/faculty/Publication%20Files/08-039_1861e507-1dc1-4602-85b8-90d71559d85b.pdf.
11. Miotto, Riccardo, Brian A. Kidd, Joel T. Dudley. "Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records," *Scientific Reports*, 17 mai 2016, https://www.nature.com/articles/srep26094.
12. Mordvintsev, Alexander, Christopher Olan, Mike Tyka. "Inceptionism: Going Deeper into Neural Networks," *Google AI Blog*, June 17, 2015, https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html.
13. "Inceptionism: Going deeper into Neural Networks," *Google Photos*, 17 iunie 2015, https://photos.google.com/share/AF1QipPX0SCl7Oz-Wilt9LnuQliattX4OUCj_8EP65_cTVnBmS1jnYgsGQAieQUc1VQWdgQ?key=aVBxWjhwSzg2RjJWLWRuVFBBZEN1d205bUdEMnhB.
14. Sweeney, Latanya."Discrimination in Online Ad Delivery," *Harvard University*, 28 ianuarie 2013, https://arxiv.org/abs/1301.6822.
15. Winston, Ali. "Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology," *The Verge*, 27 februarie 2018, https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd.
16. Angwin, Julia, Jeff Larson, Surya Mattu, Lauren Kirchner. "Machine Bias," *ProPublica*, 23 mai 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
17. McLaughlin, Kevin, Jessica E. Lessin. "Deep Confusion: Tensions Lingered Within Google Over DeepMind," *The Information*, 19 aprilie 2018,

18. Vincent, James. "Google's DeepMind and UK hospitals made illegal deal for health data, says watchdog," *The Verge*, 3 iulie 2017, https://www.theverge.com/2017/7/3/15900670/google-deepmind-royal-free-2015-data-deal-ico-ruling-illegal.
19. Suleyman, Mustafa, Dominic King."The Information Commissioner, the Royal Free, and what we've learned," *DeepMind* blog, 3 iulie 2017, https://deepmind.com/blog/ico-royal-free/.
20. "Microsoft Launches Fifth Generation of Popular AI Xiaoice," *Microsoft News Center*, https://www.microsoft.com/en-us/ard/news/newsinfo.aspx?newsid=article_2017091.
21. Kleeman, Sophie."Here Are the Microsoft Twitter Bot's Craziest Racist Rants," *Gizmodo*, 24 martie 2016, https://gizmodo.com/here-are-the-microsoft-twitter-bot-s-craziest-racist-ra-1766820160.
22. Lee, Peter. "Learning from Tay's introduction, 25 martie 2016, *Microsoft Official Blog*, https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/.
23. Harding, Verity and Sean Legassick. "Why we launched DeepMind Ethics & Society," *DeepMind blog*, https://deepmind.com/blog/why-we-launched-deepmind-ethics-society/.
24. "Chinese hate us, Baidu CEO admits as he vows 'people before profit,' *Reuters*, 10 mai 2016, https://www.cnbc.com/2016/05/10/baidu-ceo-tells-staff-to-put-values-before-profit-after-cancer-death-scandal.html.

CAPITOLUL 4 DE AICI LA SUPER-INTELIGENȚA ARTIFICIALĂ: SEMNALELE DE AVERTIZARE

1. Am modelat scenariile din partea a doua folosind studii din diferite surse. Referințele se găsesc în bibliografie. De asemenea, am vizitat expoziția *Robots* de la Science Museum (Londra) care explorează ultimii 500 de ani de roboți umanoizi, un loc minunat în care am putut explora temele prezentate în capitolele 5-7.
2. Floorwalker, Mike. "10 Deadly Disasters We Should Have Seen Coming," *Listverse*, 2 martie 2013, https://listverse.com/2013/03/02/10-deadly-disasters-we-should-have-seen-coming/. Vezi și Teather. David. "90-second nightmare of shuttle crew," *The Guardian*, 6 fe-

bruarie 2003, https://www.theguardian.com/world/2003/feb/06/columbia.science.
3. Brooker, Katrina. "I Was Devastated: Tim Berners-Lee, the Man Who Created to World Wide Web, Has Some Regrets," *Vanity Fair*, august 2018, https://www.vanityfair.com/news/2018/07/the-man-who-created-the-world-wide-web-has-some-regrets.
4. Berners-Lee, Tim. "The web is under threat. Join us and fight for it.," *World Wide Web Foundation*, 12 martie 2018, https://webfoundation.org/2018/03/web-birthday-29/.
5. "Subscriber share held by smartphone operating systems in the United States from 2012 to 2018," *Statista*, https://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-unit ed-states/.
6. "Primary e-mail providers according to consumers in the United States as of 2016, by age group," *Statista*, https://www.statista.com/statistics/547531/e-mail-provider-ranking-consumer-usa-age/.
7. Fernandez, Marisa. "Amazon leaves retail competitors in the dust, claims 50% of U.S. e-commerce market," *Axios*, 13 iulie 2018, https://www.axios.com/amazon-now-has-nearly-50-of-the-us-e-commerce-market-1531510098-8529045a-508d-46d6-861f-1d0c2c4a04b4.html.
8. Kleiner, Art. "The Man Who Saw the Future," *Strategy+Business*, 12 februarie 2003, https://www.strategy-business.com/article/8220?gko=0d07f.
9. Cass R. Sunstein, "Probability Neglect: Emotions, Worst Cases, and Law," Chicago Unbound, John M. Olin Program in Law and Economics Working Paper No. 138, 2001.
10. "Quick Facts 2015," *National Highway Traffic Safety Administration*, https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812348.
11. "Aviation Statistics," *National Transportation Safety Board*, https://www.ntsb.gov/investigations/data/Pages/aviation_stats.aspx.
12. Brooks, Frederick P. *The Mythical Man Month: Essays on Software Engineering, Anniversary Edition.* Addison Wesley; 2d edition, 2 august 1995.
13. Wilby, Peter. "Beyond the Flynn effect: new myths about race, family

and IQ?" *The Guardian*, 27 septembrie 2016, https://www.theguardian.com/education/2016/sep/27/james-flynn-race-iq-myths-does-your-family-make-you-smarter.
14. Condon, Stephanie. "US once again boasts the world's fastest supercomputer," *ZDNet*, 8 iunie 2018, https://www.zdnet.com/article/us-once-again-boasts-the-worlds-fastest-supercomputer/.
15. Viegas, Jen. "Comparison of Primate Brains Reveals Why Humans Are Unique," *Seeker*, 23 noiembrie 2017, https://www.seeker.com/health/mind/comparison-of-primate-brains-reveals-why-humans-are-unique.
16. Bostrom, Nick. "Ethical Issues in Advanced Artificial Intelligence," *nickbostrom.com*, 2003, https://nickbostrom.com/ethics/ai.html.
17. I. J. Good, "Speculations Concerning the First Ultraintelligent Machine," *Advances in Computers* 6_ (1965): 31–88.
18. Pratt, Gill A. "Is a Cambrian Explosion Coming for Robotics?" *American Economic Association*, vara 2015, https://www.aeaweb.org/articles?id=10.1257/jep.29.3.51.

CAPITOLUL 6 CUM SĂ ÎNVĂȚĂM SĂ TRĂIM CU MILIOANE DE TĂIETURI CU HÂRTIA: SCENARIUL PRAGMATIC

1. Ross, Casey. "IBM Watson Health hampered by internal rivalries and disorganization, former employees say," *STAT*, 14 iunie 2018, https://www.statnews.com/2018/06/14/ibm-watson-health-rivalries-disorganization/.
2. Idem.
3. Elsayed, Gamaleldin, Ian Goodfellow and Jascha Sohl Dickson. "Adversarial Reprogramming of Neural Networks," preprint edition accessed at https://arxiv.org/pdf/1806.11146.pdf.
4. Wang, Orange. "Chinese mobile payment giants Alipay, Tenpay fined US$88,000 for breaking foreign exchange rules," *South China Morning Post*, 25 iulie 2018, https://www.scmp.com/news/china/economy/article/2156858/chinese-mobile-payment-giants-alipay-tenpay-fined-us88000.

CAPITOLUL 7 DINASTIA RÉNGŌNG ZHÌNÉNG: SCENARIUL CATASTROFIC

1. "China Has a Vastly Ambitious Plan to Connect the World," *The Economist,* 28 iulie 2018, https://www.economist.com/briefing/2018/07/26/china-has-a-vastly-ambitious-plan-to-connect-the-world.
2. Idem.
3. Ibidem.
4. Ibidem.
5. Londoño, Ernesto. "From a Space Station in Argentina, China Expands Its Reach in Latin America," *New York Times,* 28 iulie 2018, https://www.nytimes.com/2018/07/28/world/americas/china-latin-america.html.
6. Kochanek, Kenneth D, Sherry L. Murphy, Jiaquan Xu, Elizabeth Areas. "Mortality in the United States, 2016," *NCHS Data Brief,* https://www.cdc.gov/nchs/data/databriefs/db293.pdf

CAPITOLUL 8 PIETRICELE ȘI BOLOVANI: CUM SĂ REPARĂM VIITORUL INTELIGENȚEI ARTIFICIALE

1. "Vinton G. Cerf," *Google AI,* https://ai.google/research/people/author32412.
2. Povestirea Runaround (Fuga în cerc) a fost publicată pentru prima dată în numărul din martie 1942 al revistei *Astounding Science Fiction.* Ea face parte și din colecțiile de povestiri *I, Robot (1950), The Complete Robot (1982) și Robot Visions (1990).*
3. "Human Cell Atlas," https://www.humancellatlas.org/learn-more.
4. Metz, Cade. "As China Marches Forward on A.I., the White House Is Silent," *New York Times,* 12 februarie 2018, https://www.nytimes.com/2018/02/12/technology/china-trump-artificial-intelligence.html.
5. Heisler, Yoni. "Amazon in 2017 spent almost twice as much on R&D as Microsoft and Apple – combined," *BGR,* 10 aprilie 2008, https://bgr.com/2018/04/10/amazon-vs-apple-research-and-devel-

opment-2017-alphabet-google/
6. "The OTA Legacy," *Princeton University*, http://www.princeton.edu/~ota/.
7. "Dining," *Department of Defense*, http://www.whs.mil/our-services/building-facilities/dining.
8. "The Spheres," https://www.seattlespheres.com/.
9. Adamczyk, Alicia. "These Are the Companies With the Best Parental Leave Policies," *Money*, 4 noiembrie 2015, http://time.com/money/4098469/paid-parental-leave-google-amazon-apple-facebook/.
10. Webb, Amy. "Apple vs. FBI debate may be the least of our challenges," *CNN*, 29 februarie 2016 https://www.cnn.com/2016/02/25/opinions/when-technology-clashes-with-law-iphone-opinion-webb/index.html.
11. "China Uncovers 500,000 Food Safety Violations in Nine Months," *Reuters*, 24 decembrie 2016, https://www.reuters.com/article/us-china-food-safety/china-uncovers-500000-food-safety-violations-in-nine-months-idUSKBN14D046.
12. Tandon, Suneera. "An IBM team identified deep gender bias from 50 years of Booker Prize shortlists," *Quartz India*, 24 iulie 2018, https://qz.com/india/1333644/ibm-identifies-gender-bias-in-booker-prize-novel-shortlists/.
13. Mason, Hilary. *Twitter*, 28 martie 2018, https://twitter.com/hmason/status/979044821749895170

În loc să pună sub lupă caracterul mașinilor gânditoare, futurista Amy Webb întoarce o privire critică asupra oamenilor din spatele computerelor. Cu o Inteligență Artificială coordonată într-o măsură copleșitoare, ce posibilitate există ca tehnologia să fie pusă în slujba interesului tuturor oamenilor, se întreabă ea.

-Wired

Webb reușește să picteze un tablou viu al modului în care IA ar putea lucra pentru binele omului obișnuit, de la medicina de precizie până la aplicațiile de dating...Predicțiile ei sunt provocatoare și neliniștitor de plauzibile.

-Science News

În loc să prezică viitorul, Webb creionează scenarii cu consecințe optimiste, pragmatice și catastrofice, toate construite pornind de la date curente. Indiferent cât de imposibilă vi s-ar părea ideea unui sistem de operare comun Apple-Amazon, Applezon, analizarea unor scenarii potențiale este un exercițiu fantastic de sănătos,pentru că toți cei care vă vor spune că știu exact cum va evolua IA, pur și simplu mint.

-VentureBeat

Trebuie s-o aducem pe Amy Webb în campus. Acesta e unul dintre acele cazuri în care organizarea unei discuții despre carte ar fi grozavă, dar ineficientă. Cel mai înțelept ar fi să interacționăm direct cu Webb... Avem nevoie de mai multe cărți precum aceasta, care critică educația academică din motive care transcend politica(prejudecata liberală) sau costurile. Este, în mod esențial, o carte pentru toți cei interesați de o perspectivă globală asupra rolului companiilor și a politicilor globale în determinarea schimbării tehnologice.

-InsideHigherEd

Webb e o povestitoare de prima mână; a petrecut o viață întreagă studiind și scriind, convertindu-și ideile în cuvinte, mai precis în această capodoperă a ei.

-Law.com

Cei nouă titani tech ne oferă perspective îndrăznețe despre viitorul inteligenței artificiale. Spre deosebire însă de alți autori care încearcă să anticipeze viitorul,Webb nu se folosește de senzaționalism, ci de argumente studiate, un context istoric profund și un nivel înfricoșător de plauzibilitate.

-Jonathan Zittrain, profesor de drept internațional și informatică, UniversitateaHarvard

#CEINOUATITANITECH

ISBN 978-606965300-5

9 786069 653005

globo EDITURA

COMANDĂ ON-LINE
www.edituraglobo.ro